

Paper 159-29

Implementation of Web-based ROLAP

Haidong Tang, Xiao Ji

Data Strategies Dept., Shanghai Baosight Software Co., Ltd., China

ABSTRACT

ROLAP (Relational OLAP) is a technology used to allow end users to quickly analyze relational stored data from different dimensions. SAS® procedure TABULATE is a powerful tool to create tabular reports, while it can be an ideal tool to develop ROLAP applications as well. By establishing hierarchical relationships among analysis variables, PROC TABULATE can realize ROLAP operations as drill-down, drill-up, and rotation. With SAS Application Dispatcher within SAS/IntrNet® software and ODS (Output Delivery System) technology, this kind of ROLAP mechanism can be easily implemented on Web. This paper is to introduce the essence of ROLAP and the implementation of Web-based ROLAP using SAS.

INTRODUCTION

OLAP (Online Analysis Processing) is a technology used to quickly analyze data from different dimensions. According to different physical storage, OLAP has been classified into MOLAP (Multi-dimensional OLAP), ROLAP (Relational OLAP) and HOLAP (Hybrid OLAP). Though these three types of OLAP have different aspects, ROLAP is the most convenient to be acquainted, and is the most adaptable to average end users. A ROLAP application is to interactively analyze the statistics variables via different analysis variables and their hierarchical combinations, which are also known as dimensions.

The TABULATE procedure is a powerful method to display descriptive statistics in tabular format. PROC TABULATE provides flexibility in classifying the values of variables and establishing hierarchical relationships between the variables. SAS Application Dispatcher and ODS are useful technologies to develop applications on Web. We are going to discuss the essence of ROLAP using the TABULATE procedure, and its implementation on Web.

ESSENCE OF ROLAP

ROLAP is a form of online analysis processing (OLAP) that performs dynamic multidimensional analysis of data stored in a relational database, rather than in a multidimensional database (or cube, which is usually considered as OLAP standard). Typically there are two kinds of variable in a relational table, analysis variables and statistics variables. Analysis variables are usually known as characters, while statistics variables as numbers. ROLAP is to interactively show statistics summary according to different analysis variables and their hierarchical combinations. Drill-down, drill-up, and rotation are actually based on the combination and hierarchy of all different analysis variables.

To make it simple and clear, we will discuss it further with an example of a single relational SAS table. Let's take the data set sashelp.class as an example. In order to interactively analyze the relationships among analysis variables SEX and AGE, and statistics variable WEIGHT, we can first build up the following summary report,

	Weight
	Sum
F	811.00
M	1089.50

Figure 1

And then we can drill-down from SEX by AGE,

		Weight
		Sum
F	11	50.50
	12	161.50
	13	182.00
	14	192.50
	15	224.50
M	11	85.00
	12	310.50
	13	84.00
	14	215.00
	15	245.00
	16	150.00

Figure 2

The output from Figure 1 to Figure 2 is a typical drill-down operation. (Here AGE has been taken as an analysis variable, though it is actually a numerical variable in the table.)

We can also rotate the above tabular report,

F					M					
11	12	13	14	15	11	12	13	14	15	16
Weight										
Sum										
50.50	161.50	182.00	192.50	224.50	85.00	310.50	84.00	215.00	245.00	150.00

Figure 3

An analysis variable can be a dimension, while different combination and different hierarchy of analysis variables can result to different aspects of statistics. So we can see that a ROLAP is essentially a tabular report with interactively operations of drill-down, drill-up and rotation according to the combination of analysis variables.

BASIC ROLAP IMPLEMENTATION

PROC TABULATE is a powerful and flexible tool to produce various tabular reports according to different analysis variables and their combinations. All ROLAP operations can be simulated by the procedure, and it is actually an ideal tool to implement ROLAP applications.

Here is an example to stimulate drill-down using PROC TABULATE. Following code generates Figure 1,

```
proc tabulate data=sashelp.class;
  class age sex;
  var weight;
  table sex='',weight;
```

```
run;
```

And the code below generates Figure 2, which is likely a drill-down operation.

```
proc tabulate data=sashelp.class;
  class age sex;
  var weight;
  table sex=' '*age=' ',weight;
run;
```

Following code is to rotate the tabular report from Figure 2 to Figure 3.

```
proc tabulate data=sashelp.class;
  class age sex;
  var weight;
  table (sex=' '*age=' '*weight);
run;
```

To implement a drill-down, we should only pass analysis variable AGE to the second code script. And to implement a rotation, we should only rearrange codes in TABLE option. The problem is that how we can make the process smooth and easy to operate. And SAS Web technologies provide the answers.

WEB-BASED ROLAP

SAS Application Dispatcher available within SAS/IntrNet software has proved to be an efficient and flexible Web authoring tool. Its powerful functions can help to fulfill the ROLAP applications based on Web. These applications can transfer parameters from one session to another. This solves quite well the problem of passing analysis variables from one operation to another, which might be a drill-down, or a drill-up, or a rotation operation.

We are to take the following relational table as an example to discuss further.

```
Actual_Sales num format=COMMA12. informat=COMMA32.,
Country char(12) format=$12. informat=$12.,
Region char(12) format=$12. informat=$12.,
Division char(12) format=$12. informat=$12.,
Quarter num format=BEST12. informat=BEST32.,
Year num format=BEST12. informat=BEST32.,
Month char(3) format=$3. informat=$3.
```

To implement a Web-based ROLAP application, first we should distinguish all the analysis variables (dimensions) and statistics variables. Here is the sample page of variables selection. Listed in "Analytical Vars" are all analysis variables, which mostly are characteristic variables. And in "Statistical Vars" are all statistics variables available.

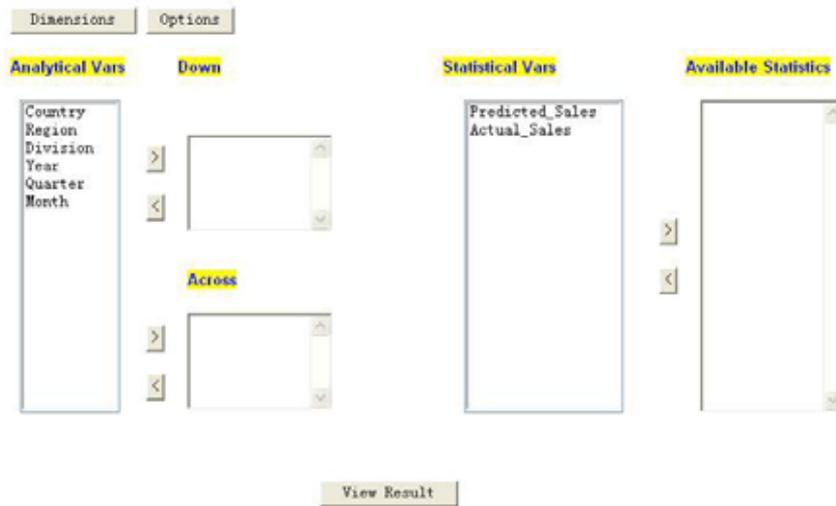


Figure 4

We should then decide the root level analysis variables, both vertically and horizontally. We choose Country and Year as vertical and horizontal variables respectively. And for each statistics variable, there are several statistics such MEAN, MAX, MIN, SUM, PCTSUM, etc. We choose two statistics “Actual_Sales: SUM” and “Actual_Sales: PCTSUM” from “Available Statistics”.

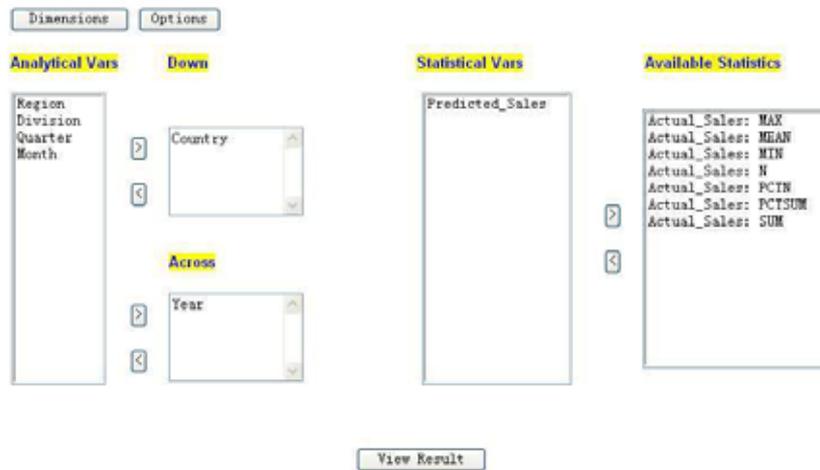


Figure 5

Here comes the initial layout of the ROLAP tabular report.

	Year ▾			
	1993		1994	
	Sum	PctSum	Sum	PctSum
Country ▾				
CANADA	121020.00	16.57	125970.00	17.25
GERMANY	127404.00	17.44	118594.00	16.24
U.S.A.	121053.00	16.57	116296.00	15.92

Figure 6

There are two drillable variables, Country and Year. Moving mouse on the arrow right to Country, we can see all drill-down variables.

	Year ▾			
	1993		1994	
	Sum	PctSum	Sum	PctSum
Country ▾	Region			
	Division			
	Quarter			
	Month			
CANADA	121020.00	16.57	125970.00	17.25
GERMANY	127404.00	17.44	118594.00	16.24
U.S.A.	121053.00	16.57	116296.00	15.92

Figure 7

Drill-down from Country by Region, we can get the following tabular report.

		Year ▾ ▾			
		1993		1994	
Country ▾	Region ▾ ▾				
CANADA	EAST	64130.00	8.78	63355.00	8.67
	WEST	56890.00	7.79	62615.00	8.57
GERMANY	EAST	63172.00	8.65	61375.00	8.40
	WEST	64232.00	8.79	57219.00	7.83
U.S.A.	EAST	60227.00	8.25	58002.00	7.94
	WEST	60826.00	8.33	58294.00	7.98

Figure 8

Also, rotation can be performed easily too.

	Country											
	CANADA				GERMANY				U.S.A.			
	Region				Region				Region			
	EAST		WEST		EAST		WEST		EAST		WEST	
	Sum	PctSum										
Year												
1993	64130.00	8.78	56890.00	7.79	63172.00	8.65	64232.00	8.79	60227.00	8.25	60626.00	8.33
1994	63365.00	8.67	62615.00	8.57	61375.00	8.40	57219.00	7.83	58002.00	7.94	58294.00	7.98

Figure 9

The tricky thing is the way to pass the all reasonable analysis variables for drill-down and drill-up from session to session. Therefore, we should relate all reasonable analysis variables to the clickable up-and-down arrows right to analysis variables.

ADVANTAGES OF WEB-BASED ROLAP

The Web-based ROLAP mechanism introduced here are simple, and easy to implement using SAS. When we are to develop such a ROLAP application, we should never think about the details of statistics calculation. We should just concentrate on the ways to pass the analysis and statistics variables to PROC TABULATE, and fit them well to the option TABLE, so as to simulate operations of drill-down, drill-up and rotation.

As ROLAP is based on a single table, it is usually easy to be adopted by end users. It is common that ROLAP is related to Star-Schema data modeling, which is often used in data warehouse system. So this kind of ROLAP mechanism can work well with data warehouse system.

For end users of Web-based ROLAP, only a Web browser is needed on their PCs. Therefore it is convenient for them to operate, and often easy to be accepted.

CONCLUSIONS

The mechanism introduced in this paper takes the most advantage of PROC TABULATE and SAS Application Dispatcher. As the real situation may be varied, the implementation of ROLAP applications can be a little bit varied. However this method can be used to develop a custom ROLAP tool, which can be easily deployed with SAS data warehouse solution.

REFERENCES

SAS Online Document, V8 - SAS Institute Inc., Cary, NC

CONTACT INFORMATION

Both authors can be contacted at:
 Haidong Tang, Xiao Ji,
 Data Strategies Dept.,
 Shanghai Baosight Software Corp., Ltd.
 515 Guoshoujing Rd., Pudong District,
 Shanghai 201203
 P.R.China

Email: tanghaidong@baosight.com
jixiao@baosight.com

Web: www.baosight.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.