Paper 025-2012

Managing Multiple Languages with a Single Cube using SAS® 9.3

Amela (Ying) Jin, Tony Zhang, SAS Institute Inc., Beijing, China

ABSTRACT

Global organizations often manage business analysis efforts across a worldwide corporate enterprise. When OLAP technology is used to analyze organizational data, multiple OLAP cubes that contain different languages and cultural conventions are often utilized. Managing these cubes can be costly and resource intensive for an organization. It is now possible to manage multiple languages and conventions within a single cube using SAS Unicode Server.

The example in this paper shows how source data from the United States, France, and China can be integrated. Using SAS Unicode Server, multiple languages are combined into a single cube. This provides reusable reporting that can be adapted to meet region specific needs. While multiple languages are often utilized in global organizations, those languages and conventions can now be combined into single cubes. This approach provides a united business analysis.

INTRODUCTION

In SAS 9.3, Multiple Language Support (MLS) technology is designed to manage multiple languages and conventions in a single cube. An MLS structured cube can meet cultural conventions for global users, performing business analysis based on a single OLAP cube. MLS enables an organization to reduce the number of OLAP cubes needed to maintain to region specific requirements, while simultaneously reducing maintenance cost. The example shown here demonstrates the steps needed to build an MLS Cube that supports three different languages: English, French, and Simplified Chinese. Here are the main contents:

- Preparing data of three languages

- Designing and creating a cube using SAS OLAP Cube Studio

- Creating a cube using Proc OLAP

- Viewing the cube in different clients: SAS Add-in for Microsoft Office, SAS Enterprise Guide and SAS Web Report Studio

In clients with different language settings, the MLS cube displays multilingual data as shown with the following figure.
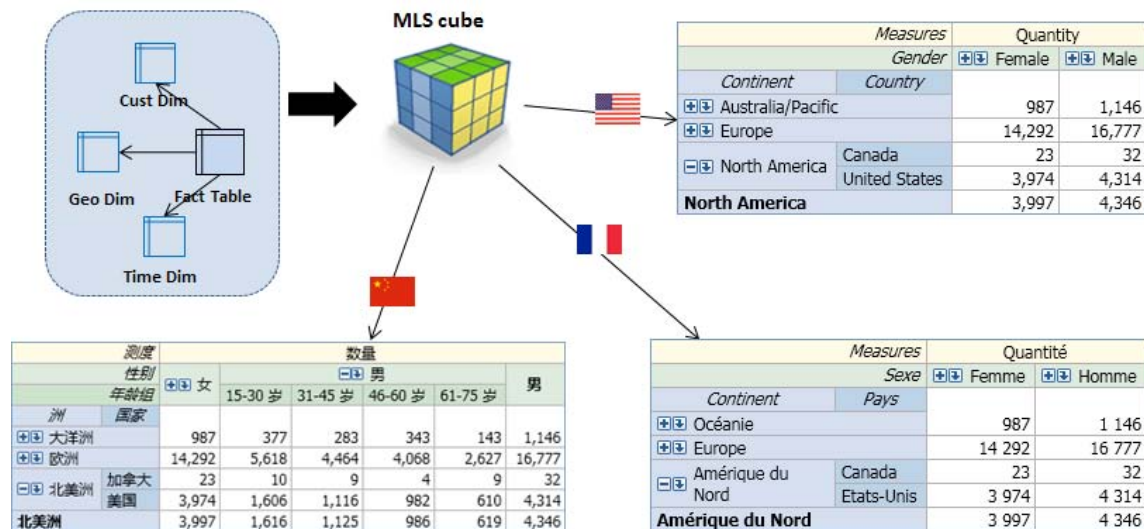

Figure 1. Multiple Language Support Cube

CONFIGURING A SAS UNICODE SERVER

Before addressing the details of data preparation and cube design, you must make sure the SAS environment is capable of handling data of different languages.

Managing Multiple Languages with a Single Cube using SAS® 9.3, continued

For example, French and Chinese customer data can be processed in SAS French and Chinese sessions separately. The encoding of the SAS French session is wlatin1 and the Chinese session is euc-cn. To read and write French and Chinese data in a single SAS session, a universal encoding that can handle multilingual characters is needed. With the SAS Unicode server, whose encoding is UTF-8, it is possible to store, process, and deliver multilingual data.

In SAS 9.3, the SAS Deployment Wizard has an option to configure a Unicode server.
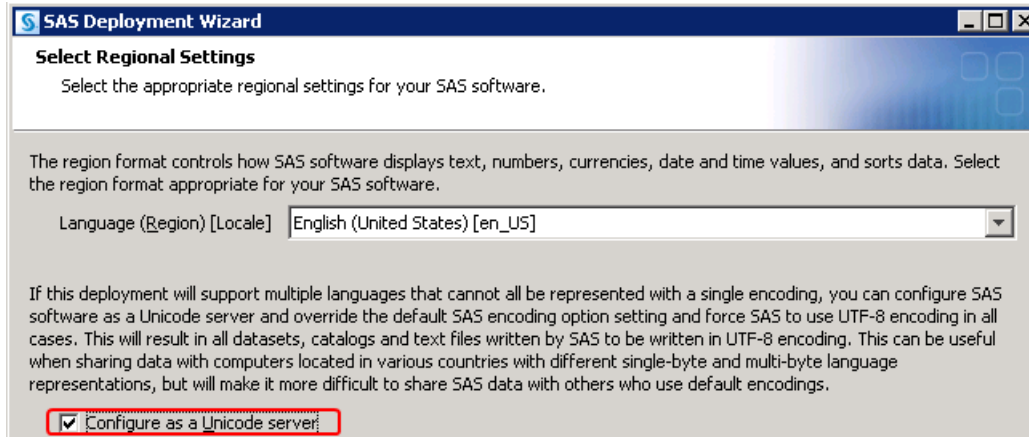
Figure 2. SAS Deployment Wizard Regional Settings

Note: If you have an existing non-Unicode SAS environment, you can also change the configurations. Please see: *Best Practices for a Non-English BI Deployment.*

PREPARING DATA

The input data used in this example is sales data of a fictitious global sports and outdoors retailer named Orion Star. The corporate headquarters are located in the United States. Orion Star also has offices and stores in many countries around the world.

This company collects following data:

- Customer age group and gender

- Sales location

- Order quantity, date, and related information

A star schema is used for input in this example. The data for the cube is represented in the following figure:
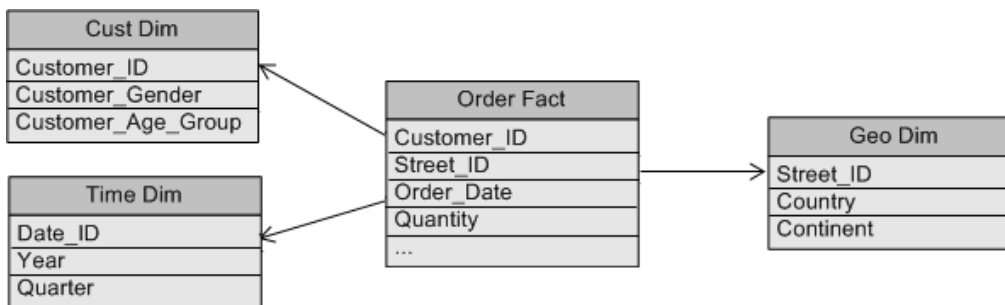
Figure 3. Cube Input Data

Unlike a single language SAS OLAP cube, an MLS cube requires additional translation tables including dimension table, caption table, and drill-through table.

In this example, an OLAP cube is designed to support analysis in English, French, and Simplified Chinese. Translated dimension tables, caption tables, and drill-through tables are prepared in three different languages.

Translated Dimension Tables

Dimension tables share a common base name. The base name identifies the dimension table for the default

Managing Multiple Languages with a Single Cube using SAS® 9.3, continued

language. Dimension table names for the other languages are composed of the base name plus the Posix locale name. For each translated dimension, you must create one translated dimension table for each locale that you plan to support. The MLS_GEO_DIM table contains English geography information. The translated copies of this dimension table are MLS_GEO_DIMfr_fr for French (France) and MLS_GEO_DIMzh_cn for Chinese (China).

Figure 4. Translated GEO Dimension Tables

Note: A multilingual cube can support any of the locales that are valid for the SAS system option LOCALE=. You can find a list for the values of LOCALE= system option in *SAS 9.3 National Language Support (NLS): Reference Guide*.

Translated Caption Tables

Translated caption tables share the same naming requirements as translated dimension tables. You need at least one translated caption table for each locale that you plan to support.

Within your translated caption tables, captions are referenced by a numeric MLSID. Any translated copies of the default caption table would apply the same MLSID to the same translated caption, as shown in following example:

Figure 5. Translated Caption Tables

MLSID values can be assigned to any cube object that can receive a caption. You can use one set of caption tables for all captions in cube, or use separate tables for cube measures and each dimension. In this example, the cube data uses one set of caption tables for all captions in cube.

Translated Drill-through Table

Translated drill-through tables share the same naming requirements as translated dimension tables. You need to create one translated drill-through table for each language you plan to support.

Figure 6. Translated Drill-through Tables

Here's a list of all the translated tables used in this example:

Table Type	English (Default)	French	Simplified Chinese
Translated Dimension Table	MLS_CUST_DIM	MLS_CUST_DIMFR_FR	MLS_CUST_DIMZH_CN
	MLS_GEO_DIM	MLS_GEO_DIMFR_FR	MLS_GEO_DIMZH_CN
	MLS_TIME_DIM	MLS_TIME_DIMFR_FR	MLS_TIME_DIMZH_CN

Managing Multiple Languages with a Single Cube using SAS® 9.3, continued

Table Type	English (Default)	French	Simplified Chinese
Translated Caption Table	CAP	CAPFR_FR	CAPZH_CN
Translated Drill-Through Table	ORDER_DETAIL	ORDER_DETAILFR_FR	ORDER_DETAILZH_CN

**Table 1. Translated Tables List**

CREATING THE MLS CUBE

This example shows two methods to create an MLS cube.

1. Create the cube using SAS OLAP Cube Studio,

2. Create the cube using the SAS OLAP procedure.

SAS OLAP Cube Studio provides a Cube Designer wizard that guides you through steps to create and modify a cube. However, if you prefer to define a cube in batch mode, you can define and build SAS OLAP cubes using the OLAP procedure.

Creating a Cube Using SAS OLAP Cube Studio

1. In SAS OLAP Cube Studio, select **File - New - Cube**. In the **Cube Designer - General** window, specify your cube name, location, and cube path. Multiple language support is available only for cubes that are loaded from star schemas. You must select **Star schema** as **Input Type**.
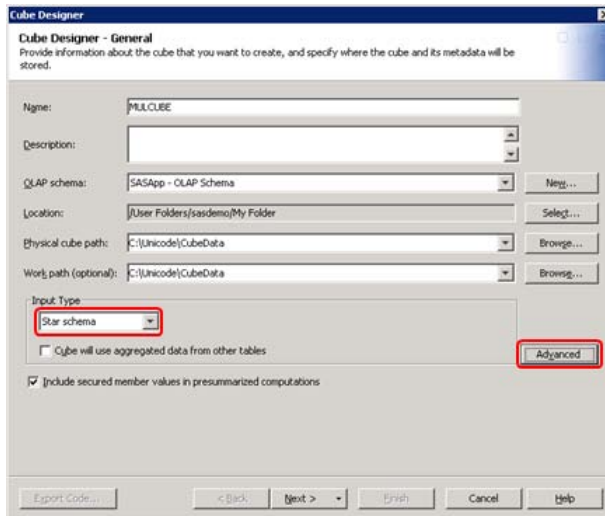
Figure 7. Cube Designer - General

2. Click the **Advanced** button on the **Cube Desinger - General** window. In the **Advanced Cube Options** dialog box, select the **Dimension Table Translation** tab. Select languages and locales for the cube. In this example, the first selection is the default language and locale of your cube.

Managing Multiple Languages with a Single Cube using SAS® 9.3, continued
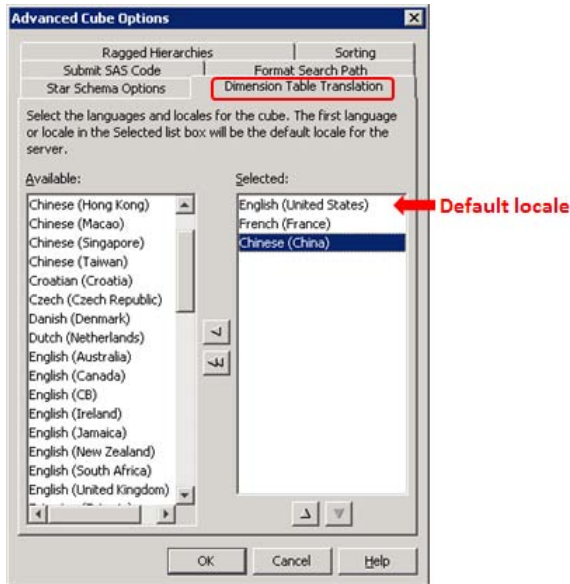
Figure 8. Advanced Cube Option

3. In the **Cube Designer – Input** window, select fact table and drill-through table. Select the **Support drill-through tables in multiple languages** option.
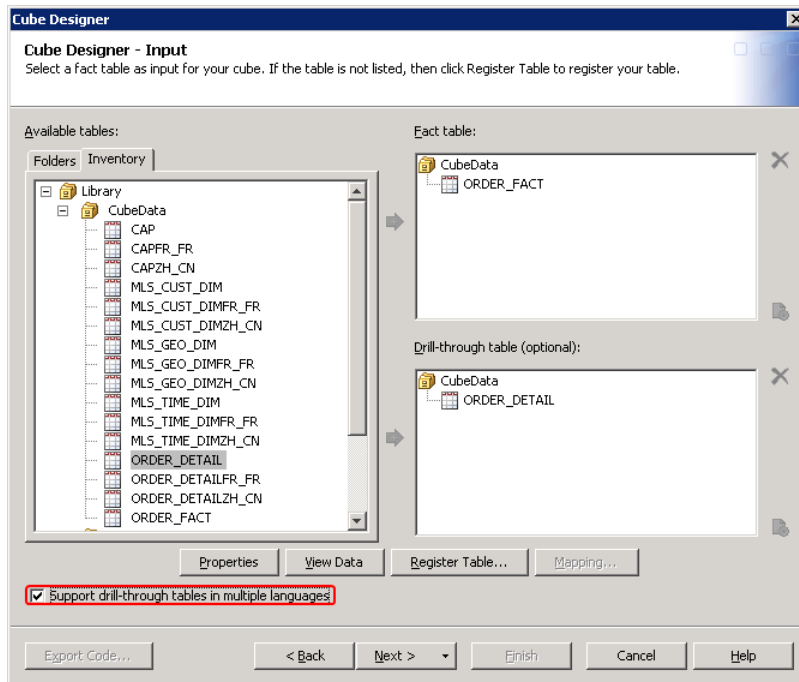
Figure 9. Cube Designer – Input

4. In the **Cube Designer – Dimension Tables** window, select MLS_CUST_DIM, MLS_GEO_DIM, and MLS_TIME_DIM tables.

5. In the **Cube Designer – Dimensions** window, add levels and hierarchies for the three dimensions.

6. In the **Cube Designer – Select Measures** window, select cube measures.

7. In the **Cube Designer – Aggregations** window, add aggregations that you need.

8. In the **Cube Designer – Caption Table Translations** window, select the **Optionally, specify tables that contain the translated captions for the cube and its components** option. Select your cube caption table.

Managing Multiple Languages with a Single Cube using SAS® 9.3, continued
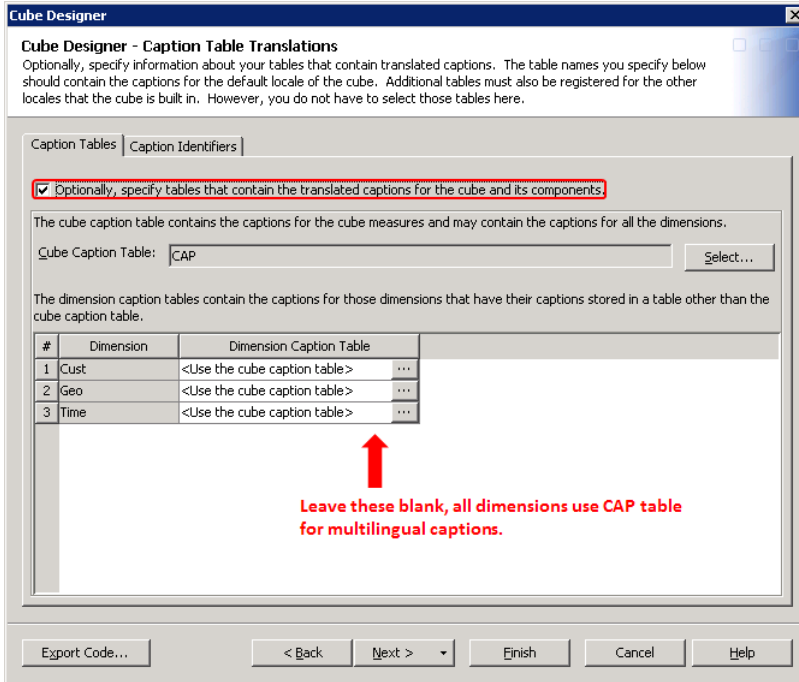
Figure 10. Cube Designer - Caption Tables

9.   On the **Caption Identifier** tab, enter identifiers for the specific components.
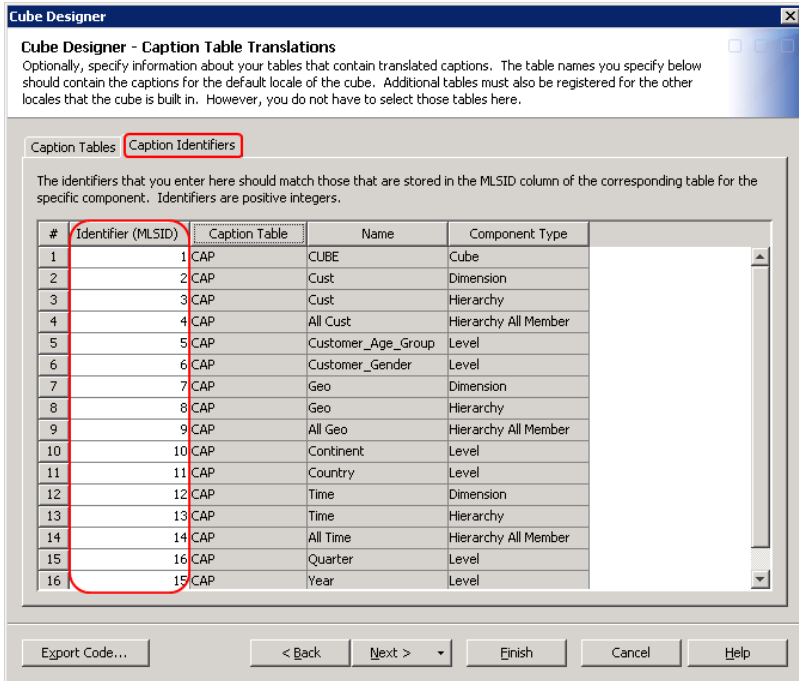
Figure 11. Cube Designer - Caption Identifiers

The identifiers should match values in MLSID column of your caption table.

10. On the **Summary of selections** window, review the settings for the cube, and then click **Finish** to display the Finish dialog box. You can now build the cube.

Creating a Cube Using OLAP Procedure

To create an MLS cube, there are several important language elements you should be aware of in OLAP procedure:

1.   USER_DEFINED_TRANSLATION statement

The USER_DEFINED_TRANSLATION statement specifies the languages and locales your cube supports, and thus it implicitly defines your translated table names.

2.   Options in PROC OLAP statement

DTLIBREF specifies a library for translated drill-through tables.

DTMEMPREF specifies the prefix of the drill-through table names.

CUBETABLELIBREF specifies a library for translated caption tables.

CUBETABLECAPPREF specifies the prefix of the translated caption table names.

3.   Options in DIMENSION statement

DIMTABLELIBREF specifies a library for translated dimension and/or caption tables.

DIMTABLEMEMPREF specifies the prefix of the translated dimension table names.

DIMTABLECAPPREF specifies the prefix of the translated caption table names.

4.   MLSID option

For any HIERARCHY, LEVEL, PROPERTY, and MEASURE statement that provides a caption, you can use MLSID option to identify the row in the caption tables that contains the translated captions.

Here is an example of PROC OLAPcode. Multiple language support elements are marked in bold.

```
OPTIONS VALIDVARNAME=ANY;
LIBNAME DEMO BASE "C:\demo\data";

PROC OLAP CUBE="/Shared Folders/MLSCUBE" PATH='c:\demo\cube' FACT=DEMO.ORDER_FACT
    CUBETABLELIBREF=DEMO          /*library for translated caption tables*/
    CUBETABLECAPPREF=CAP          /*prefix of translated caption table names*/
    DTLIBREF=DEMO                 /*library for translated drill-through tables*/
    DTMEMPREF=ORDER_DETAIL        /*prefix of translated drill-through table names*/
    MLSID=1;

METASVR HOST="servername" PORT=8561 OLAP_SCHEMA="SASApp - OLAP Schema"
userid="xxxx" pw="xxxx";

DIMENSION Cust SORT_ORDER=ASCENDING DIMKEY=Customer_ID FACTKEY=Customer_ID
HIERARCHIES=(Cust)
    DIMTABLELIBREF=DEMO                  /*library for translated dimension tables*/
    DIMTABLEMEMPREF=MLS_CUST_DIM         /*prefix of translated cust dimension tables*/
    MLSID=2;                             /*MLSID of cust dimension caption*/
HIERARCHY Cust LEVELS=(Customer_Gender Customer_Age_Group) DEFAULT
    ALL_MLSID=4               /*MSLID of all member caption in cust hierarchy*/
    MLSID=3;
LEVEL Customer_Age_Group SORT_ORDER=ASCENDING
    MLSID=5;
LEVEL Customer_Gender SORT_ORDER=ASCENDING
    MLSID=6;

DIMENSION Geo TYPE=GEO SORT_ORDER=ASCFORMATTED DIMKEY=Street_ID FACTKEY=Street_ID
HIERARCHIES=(Geo)
    DIMTABLELIBREF=DEMO
    DIMTABLEMEMPREF=MLS_GEO_DIM
    MLSID=7;
HIERARCHY Geo LEVELS=(Continent Country) DEFAULT
    ALL_MLSID=9
    MLSID=8;
LEVEL Continent SORT_ORDER=ASCFORMATTED
    MLSID=10;
LEVEL Country SORT_ORDER=ASCFORMATTED
    MLSID=11;
```

Managing Multiple Languages with a Single Cube using SAS® 9.3, continued

```
DIMENSION Time TYPE=TIME SORT_ORDER=ASCENDING DIMKEY=Date_ID FACTKEY=Order_Date
HIERARCHIES=(Time)
    DIMTABLELIBREF=DEMO
    DIMTABLEMEMPREF=MLS_TIME_DIM
    MLSID=12;
HIERARCHY Time LEVELS=(Year Quarter) DEFAULT
    ALL_MLSID=14
    MLSID=13;
LEVEL Year COLUMN=YEAR_ID TYPE=YEAR SORT_ORDER=ASCFORMATTED
    MLSID=15;
LEVEL Quarter COLUMN=QUARTER TYPE=QUARTERS SORT_ORDER=ASCFORMATTED
    MLSID=16;

MEASURE QuantitySUM STAT=SUM COLUMN=Quantity FORMAT=NLNUM15. DEFAULT
    MLSID=17;

/*locales supported*/
USER_DEFINED_TRANSLATIONS EN_US ZH_CN FR_FR;

RUN;
```

VIEWING A CUBE USING SAS CLIENTS

You can view and query SAS OLAP cubes using SAS clients such as SAS Enterprise Guide, Microsoft Office Excel with SAS Add-In, and SAS Web Report Studio. These clients provide different levels of support for MLS cube and have different mechanism to determine what locale is currently used.

Viewing a Cube in SAS Enterprise Guide

In SAS Enterprise Guide, the dimension and caption for a cube are displayed according to the application's language settings. To check language settings of SAS Enterprise Guide, select **Options** from **Tools** menu.

In the **General** options, the **Current selected language** list displays the language currently used, such as **Chinese (People's Republic of China)**. If this setting is changed, you must restart SAS Enterprise Guide after your change.

测度		数量					
性别		⊞⊡ 女	⊟⊡ 男				男
年龄组			15-30 岁	31-45 岁	46-60 岁	61-75 岁	
洲	国家						
⊞⊡ 大洋洲		987	377	283	343	143	1,146
⊞⊡ 欧洲		14,292	5,618	4,464	4,068	2,627	16,777
⊟⊡ 北美洲	加拿大	23	10	9	4	9	32
	美国	3,974	1,606	1,116	982	610	4,314
北美洲		3,997	1,616	1,125	986	619	4,346

Figure 12. Viewing a Cube in SAS Enterprise Guide

You can drill through a cell to get detailed information in the selected language. Right click on the cell you are interested in and click **Drill Through Details**.

	订货日期	顾客性别	顾客年龄组	国家	洲	年	季度
1	2011年01月02日	女	46-60 岁	加拿大	北美洲	2011	2011Q1
2	2011年01月03日	男	31-45 岁	加拿大	北美洲	2011	2011Q1
3	2011年01月08日	男	31-45 岁	加拿大	北美洲	2011	2011Q1
4	2011年01月10日	女	15-30 岁	加拿大	北美洲	2011	2011Q1
5	2011年01月14日	男	61-75 岁	加拿大	北美洲	2011	2011Q1
6	2011年01月20日	男	15-30 岁	加拿大	北美洲	2011	2011Q1
7	2011年01月21日	男	31-45 岁	加拿大	北美洲	2011	2011Q1
8	2011年01月29日	男	46-60 岁	加拿大	北美洲	2011	2011Q1
9	2011年02月20日	女	15-30 岁	加拿大	北美洲	2011	2011Q1
10	2011年03月17日	女	61-75 岁	加拿大	北美洲	2011	2011Q1
11	2011年03月21日	女	46-60 岁	加拿大	北美洲	2011	2011Q1
12	2011年03月23日	女	15-30 岁	加拿大	北美洲	2011	2011Q1

Figure 13. Drill-through Details in SAS Enterprise Guide

If you select a language in SAS Enterprise Guide, but the cube does not support it, the default language of the cube will be used. For example, if you choose Japanese, the cube will display in English which is the default language of the cube.

Managing Multiple Languages with a Single Cube using SAS® 9.3, continued

Viewing a Cube in SAS Add-In for Microsoft Excel

If you use the Pivot Table function in Microsoft Excel to view a SAS OLAP cube, cube information displays in accordance with Microsoft Office language settings.

You can check Microsoft Office language settings with the following steps:

For Microsoft Office 2007, open **Microsoft Office Tools** - **Microsoft Office 2007 Language Settings**. On the **Display Language** tab, the **Display Microsoft Office menus and dialog boxes in:** dropdown list shows the current language used, such as **française**.

For Microsoft Office 2010, open **Microsoft Office 2010 Tools** - **Microsoft Office 2010 Language Preferences**. The default display language is listed such as **française**.

An important exception to this setting entails that if measures in your cube use international formats (for example: NLNUM*w.d)*, the formatted data display is in accordance with your client machine's Windows Regional Options. You should always use consistent language settings for your client environment, including Windows Regional and Language Options and Microsoft Office language settings.

Viewing a Cube in SAS Web Report Studio

In SAS Web Report Studio, you can open a SAS OLAP cube or use it as data source directly. SAS Web Report Studio displays a cube in the language that is in accordance with user's web browser language settings.

To check the user's Internet Explorer language settings, click **Tools** - **Internet Options** in Internet Explorer.

On the **General** tab, click the **Languages** button. The first language in the list is the language currently used, such as **English (United States) [en-us]**.

VERIFYING LANGUAGES ASSOCIATED WITH A CUBE

If you have a cube that is created by other users, you might not know what languages this cube supports.

In SAS OLAP Cube Studio or SAS Data Integration Studio, you can easily find out what languages are associated with a cube by checking the cube's properties.

On the **Structure** tab of cube properties window, the **Dimension table translation** value gives a list of languages associated with this cube.
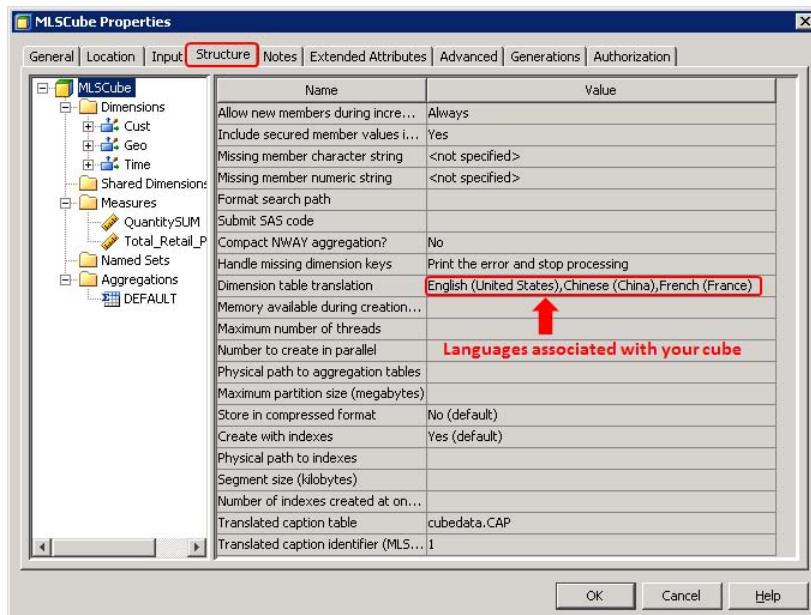
Figure 14. Cube Properties

You can check the cube metadata directly. To do this, issue the METABROWSE command in a SAS Base session and fill in the Metadata Server information. Then follow the path: Foundation - Cube - *<Your cube name>* - PropertySets - SetProperties. There is a UDT node for each language. DefaultValue attribute gives you the exact

Managing Multiple Languages with a Single Cube using SAS® 9.3, continued

locale name.

CONCLUSION

SAS Unicode Server provides the ability to store and process data in different languages. SAS MLS cube processes query of different locales and returns result accordingly. In addition, various SAS clients enable you to view an MLS cube in your language. It's especially useful for international organizations in which people have their own locale preference. You can easily build a cube for users in different regions and add more languages to your cube when needed.

TERMINOLOGY

A **LOCALE** reflects the language, local conventions such as data formatting, and culture for a geographical region. Local conventions might include specific formatting rules for dates, times, and numbers and a currency symbol for the country or region. Collating sequence, paper size, postal addresses, and telephone numbers can also be included in locale.

An **ENCODING** is the association of a unique numeric value with each symbol and punctuation mark in a character set. There are two groups or types of encodings: single byte character set (SBCS) encodings and DBCS encodings. SBCS encodings represent each character in a single byte. DBCS encodings require a varying number of bytes to represent each character. A more appropriate term for "DBCS" is multi-byte character set (MBCS). MBCS is sometimes used as a synonym for DBCS. SBCS encodings are limited to 256 possible characters. DBCS encodings can represent more than 256 characters.

UNICODE is the universal character encoding that supports the interchange, processing, and display of characters and symbols found in the world's writing systems. Other character encodings are limited to subsets of all languages. Often the subsets are limited to a region (for example: Windows Latin 1 [WLATIN1] represents the characters of the US and Western Europe on Windows). UTF8 is the Unicode encoding form where characters are represented in sequence of one to four bytes in length.

REFERENCES

SAS Institute Inc. 2011. SAS® 9.3 OLAP Server: User's Guide. Cary, NC: SAS Institute Inc.

SAS Institute Inc. 2011. SAS® 9.3 National Language Support (NLS): Reference Guide. Cary, NC: SAS Institute Inc.

SAS Institute Inc. 2010, "Best Practices for a Non-English BI Deployment." Proceedings of SAS Global Forum 2010. Cary, NC: SAS Institute Inc. Available at http://support.sas.com/resources/papers/proceedings10/336-2010.pdf.

SAS Institute Inc. 2009. "Processing Multilingual Data with the SAS® 9.2 Unicode Server." Available at http://support.sas.com/resources/papers/92unicodesrvr.pdf.

Wilkie, Michelle. "Multiple Languages Surfaced via a Cube using SAS® 9.1 OLAP Server." Available at http://support.sas.com/techsup/technote/ts701.pdf.

ACKNOWLEDGMENTS

The authors wish to express special thanks to Kansun Xia. He suggested a lot to this paper from the beginning and exerted lots of efforts to review and revise this paper. Also, we wish to thank following people (listed in alphabetical order) who have taken time to review and contribute to this paper: Holmes Niu, Jin Luo, and Xinjian Cao.

RECOMMENDED READING

SAS Institute Inc. 2011. SAS® 9.3 Intelligence Platform: Data Administration Guide. Cary, NC: SAS Institute Inc.

SAS Institute Inc. 2011. SAS® 9.3 Intelligence Platform: Security Administration Guide. Cary, NC: SAS Institute Inc.

SAS Institute Inc. 2011. SAS® 9.3 Intelligence Platform: Desktop Application Administration Guide. Cary, NC: SAS Institute Inc.

SAS Institute Inc. 2011. SAS® Web Report Studio 4.3: User's Guide. Cary, NC: SAS Institute Inc.

SAS Institute Inc. 2011. SAS® 9.3 OLAP Server: MDX Guide. Cary, NC: SAS Institute Inc.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Managing Multiple Languages with a Single Cube using SAS® 9.3, continued

Amela (Ying) Jin, Tony Zhang
SAS Research and Development (Beijing) Co., Ltd.
19/F North Wing, Central Office Tower, Junefield Plaza, No. 10 Xuanwumenwai Dajie, Xicheng District
Beijing, China
100052
Work Phone: +86 10 8319-3641/3633
Fax: +86 10 6310-9130
E-mail: amela.jin@sas.com; tao.zhang@sas.com;

Managing Multiple Languages with a Single Cube using SAS® 9.3, continued