Exploiting Key Answers from Your Data Warehouse Using SAS® Enterprise Reporter™ Software

Donna Torrence, SAS Institute Inc., Cary, North Carolina
Juli Staub Perry, SAS Institute Inc., Cary, North Carolina

Abstract

You may have the most organized data warehouse and data structures in the land. You may even have easy access to your corporate-wide data files and tables, but what good is it if business users are unable to get meaningful answers from this data? The SAS® System offers a new reporting vehicle to exploit data warehouse or enterprise-wide data all from a Windows reporting environment. The SAS® Enterprise Reporter™ software, with a similar look and feel of Microsoft Office products, allows business users to concentrate on the design of reports, rather than spending effort on creating the reports via a programming language. For the information facilitators, Enterprise Reporter™ software offers an environment for data preparation and administration. Not only does Enterprise Reporter software provide both an intuitive report building and data setup facility, it also addresses the need to administer the reporting functionality, allow quick performance in creating reports or report templates from data, and distribute the results corporate-wide in a timely fashion.

Introduction

Data Warehousing provides access to enterprise-wide data, but without an intelligent reporting tool, surfacing the data into reports can sometimes be a frustrating experience for business users. These users need data that is easily identified for analysis and bottom-line business reporting. Often, they need an administrator to be the buffer between the technical aspects of the data and to help identify the data that will be used in their daily work routine. Typically, the buffer has been a manual reporting process between the end user and the administrator. A more intelligent reporting solution is desired.

This paper will first explore a basic understanding of data warehousing and the benefits of getting data accessed and managed for business exploration. It will then introduce how the SAS Enterprise Reporter software will aid an administrator by bridging the gap between the physical location of data and the logical way that business users see the data to derive meaningful results. This will be demonstrated through a step-by-step tutorial.

Data Warehousing to Decision Making

In simple terms, a data warehouse is a container of pre-inspected data. It can contain data found throughout an entire organization that can be operational, relational, hierarchical, or external in nature. Once collected, the data is cleansed and ready for general consumption. Data administrators create technical metadata (data about data) and arrange the information into subjects. This is a step in easing the frustration of business users to gain access to the data they need.
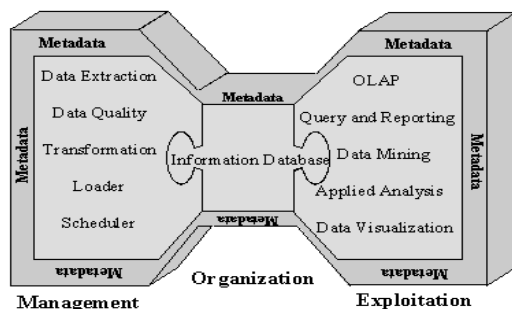
The SAS Data Warehousing Solution

Figure 1 *The SAS Data Warehousing Solution*

Figure 1 illustrates SAS® Institute's data warehousing methodology. The process can be segmented into three phases:

- Data Management
- Data Organization
- Data Exploitation

Let's explore this methodology of data warehousing in more detail.

Phase One: Data Management

The key to setting up a data warehouse is the capability and ease of use of accessing all forms of data. SAS® Institute has proven data access strategies that include:

- *MEA* (Multiple Engine Architecture) that provides a framework for translating SAS syntax for read, write and update services into RDMS' (Relational Database Management Systems) or file structures
- *Support of ANSI standard SQL* (Structured Query Language) for relational and non-relational file structures

These access methods provide a single and consistent view of enterprise data and can surface the data in two forms: as views to data or as extracts from their native form.

The SAS® System offers a variety of tools to manage this backend process. One of these tools is the SAS/Warehouse Administrator™ software. The SAS/Warehouse Administrator reduces the amount of programming work needed to build a data warehouse. It centralizes warehouse management by providing an integrated tool set, user-friendly interface, and metadata engine for generation of technical metadata about the extraction, transformation and loading processes.

Phase Two: Data Organization

Data that is stored in a form suitable for the operational systems must be integrated and transformed into subject-oriented, consistent and non-volatile data. It is necessary to perform some data management logic on the operational data, adding value to it either before or once it is placed in the data warehouse. The data structure begins to take on a business-oriented flavor, rather than applications-oriented, adding value to the data from a business perspective.

The business rules for transforming and cleaning the data are stored as metadata which are mapped to the contents of the data warehouse. This technical metadata allows the data warehouse administrators to identify the sources of data involved in a particular transformation quickly. In addition, this metadata needs to be created to support the exploration, investigation and analytical processes.

Technical metadata is information internal to an application that creates or describes elements in the application, such as tables and columns. In data warehousing, metadata can be divided into two main categories:

Physical metadata: Software statements that define how an application element stores, moves, or transforms data. For example, the physical metadata for a SAS table might specify a certain number of rows and columns, with certain data transformations applied to some columns.

Informational metadata: Text that describes the content or purpose of an application element. For example, the informational metadata for a SAS table might describe the purpose of the table and the name of the person responsible for the accuracy of the information in the table.

Phase 3: Data Exploitation

Major organizations have various types of administrators and end users. Each one requires different information from the same corporate warehouse. To support this, different metadata is needed. As mentioned

earlier, it is great for users to have access to the data needed for them to do their job. But, in order for them to make sound business decisions from this data, they need to see the data in a logical form rather than just a physical map to the data. Thus, in addition to the technical warehouse metadata, there is also a need for metadata specific to business users' reporting needs. This brings us to the final tier of the SAS data warehousing strategy: Data Exploitation.

Whether it is decision support, OLAP analysis, query and reporting or graphics, business users often measure the success of the data warehouse from getting results in these areas not from data access alone. The SAS System offers an extensive toolset for data exploitation.

SAS® Enterprise Reporter™ Software

One of these tools, the Enterprise Reporter™ software, creates, maintains and utilizes business metadata to empower business users with the ability to build meaningful, state-of-the-art reports from a data warehouse. Let's take a closer look at the Enterprise Reporter software: What is it? Who are the users? What are the various components of this tool? How will business metadata enable the business users to be more successful in getting the answers they need to drive the company forward?

What is It?

The Enterprise Reporter software is a business intelligent client/server query and reporting solution. Designed with a Microsoft Office™ look and feel, it is a PC document-centric reporting solution. Business users can produce attractive documents, based on information stored in various data structures or in a data warehouse. This reporting process empowers users with the ability to design reports before selecting data, while letting the interface drive the solution. The data is presented in familiar business terms that is transferred into columns and rows behind the scenes as it automatically generates SQL. Saved and shared reports can be accessed through the Report Gallery and/or the Intranet. All of this is possible because it utilizes the full power of the SAS System on the back-end.

Who are the Users?

Typically, a successful reporting environment within an organization will need to support a strategic hierarchy of users. For example,

- **Business Users**
  Information Consumer- many standard business users
  Information Seekers- one or more advanced business users
- **Data Administrators**
  Information Facilitators- one or few administrators

Let's investigate these users a little further.

Business User Perspective

Information Consumers are often management level and are familiar with the Microsoft Office tool set. They are users who are interested in accessing the same type of information on a daily basis. They need to examine already produced reports and produce formatted reports with graphs, summaries, and groupings that show the bottom line to help answer their business questions. In other words, they don't care about where the data resides or how it is structured, they just want to explore and report from it. They require a reporting process that supports the ability to design reports before selecting data, while letting an interface drive the solution.

Information Seekers are typically computer literate power users within various departments. They know the informational needs of their department's business area as well as the relevant data sources and how they can be used. They understand data tables and join structures and not only consume information, but also create information for others to help answer key business questions.

Data Administrator Perspective

Information Facilitators ensure business users can easily access and distribute data across the organization while still maintaining data integrity.

The Information Facilitator is familiar with a specific department's data and knows and cares about the departmental reporting needs. Very often they have access to the data store and/or data warehouse and are responsible for organizing the accessed data so the business users will be able to relate to it and report on it efficiently and accurately. This can be accomplished by providing business terms (phrases, words and terms that are recognized by their business community) through the Enterprise Reporter's metadata facility.

Specifically the Information Facilitator defines and maintains:

- users, user groups and hierarchies
- data server definitions and connections
- report data organizations
- report distribution mechanism
- customizations to the reporting environment

Some of the tasks of the facilitators can be delegated to Information Seekers (advanced users).

What are the Components of the Enterprise Reporter?

The Enterprise Reporter software creates reports that display information compiled in the organization or company. The Enterprise Reporter can be segmented into five unique interface components, all of which are tightly integrated. These five components are:

Report Builder	Environment to view, design and print reports. Other components of the reporter are accessible from this environment from the pull-down menus or toolbars. (See **Figure 10**)
Report Gallery	Holder of the standard reports and templates that are used in an organization. The Report Gallery access is controlled by the hierarchy of users and user groups defined in the User Administration component.
User Administration	Available for the system administrator to apply to groups and users based on their specific reporting need.
Server Definition	Provides the ability to point to and access remote data.
InfoCenter	Provides the ability to create and organize data unique to users' reporting needs.

Information Consumers and Seekers work most closely with the Report Builder and the Report Gallery. The Information Facilitator focuses primarily on data access, data organization and user administration. The components respective to these responsibilities are the Server Definition, the InfoCenter and the User Administration interfaces.

Now, we will take a more detailed look at data access (InfoCenter and Server Definition) components of the reporter, as this pertains specifically to utilizing data warehouse metadata.

Beginning Tutorial

We will follow a tutorial on how to access data from the corporate data warehouse or data mart, apply business rules to this data, and make it available to all types of users. We will focus on the creation of business metadata, investigate the server definition interface and finally, explore the mechanism used to manage and maintain the information needed to produce reports using the Enterprise Reporter software's InfoFolders.

Bridging the Gap Between Technical and Business Metadata

The Enterprise Reporter software metadata is data about data from a business perspective. It is designed for quick searches on warehouse structure and contains a dynamic view of the static view of data warehouse metadata. Access to this data can vary according to the location of the data

store. Some information might be stored locally on the corporate LAN (Local Area Network) or remotely, on a Unix box, for example.

Accessing Data

A first step to take is to access the data store. The SAS Enterprise Reporter software supports all industry standard databases and pc formats through the SAS System's access strategy. The reporter takes advantage of the following SAS files:

- *SAS tables*-contains columns and rows of data
- *SAS MDDB's (Multi-Dimensional Data Base)*- dimensional cubes of data at various aggregation levels
- *SAS Views*-information describing how data should be retrieved from data sources

The Enterprise Reporter software can access data on other computers/platforms. Data can be combined via file services, ODBC services and remote SAS compute services.

For local data, such as LAN and ODBC services, server definitions are not necessary. Remote data, on the other hand, which allows for additional data processing on the remote computer, requires server definitions. In the next section of this tutorial, we will investigate how to specify remote servers using the Enterprise Reporter Server Definition component.

Defining Remote SAS Servers

From within the Enterprise Reporter software, select **Tools → Server Definitions**. The Server Definition window, as shown in **Figure 2** will appear. To add a server definition, click on the **New** button in the Server box.
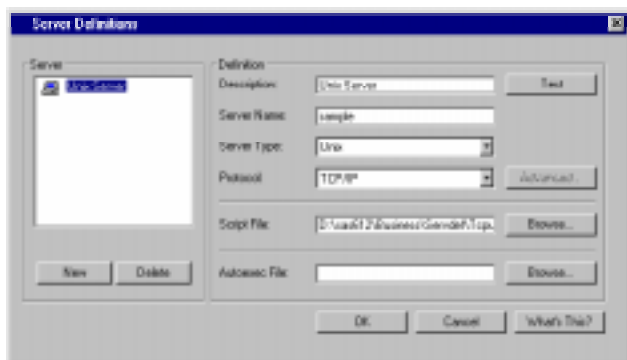

Figure 2 *Server Definition Window*

In the Server Definition window, fill in:

Description: a description to represent the remote server . Enter "Unix Server" for this example.

Server Name: the name of a named server, the IP address or the Internet address. Enter "sample " for this example.

Server Type: a valid server type. Click the down arrow to the right to get a list of valid types. Select "Unix".

Protocol: a valid protocol. Click the down arrow to the right to get a list of valid protocols. Select "TCP/IP" for this example.

Script File: a file that performs a logon to the remote system and initiates a SAS session on that server. Click on **Browse** to locate the script file named "Tcpunix.scr."

Autoexec File: a local file that is remotely executed when the connection is made. This can be used to allocate libraries that may be needed by the reporter.

Click **OK** when finished, or click on the **Test** button to test whether the connection specified can be made.   Note, when the end user selects to view the report in "Data View" of the Report Builder, the actual physical connection will be established at that time. Once remote computer locations have been specified using the method described above, point to the data sources from which the users need to report. This is accomplished by creating and, if necessary, customizing InfoFolders through the Enterprise Reporter's InfoCenter component.

InfoFolders

An InfoFolder is managed from within an InfoCenter. InfoFolders contain data items that can be selected for a report. The data items represent various forms of data that have been loaded from local or remote databases. This data can be predefined in a way that applies to the reporting needs of a business organization. For example, an InfoFolder can hold the sales information of the company. The Enterprise Reporter can contain several InfoFolders that are predefined for the users.

InfoFolders can be organized in a custom folder structure representing the physical data structure or a logical data structure or both. InfoFolders will need to be organized so that it is relevant data for departmental users and in a way that makes business sense to them. They should not be concerned with the actual physical file structures of their warehouse.

Managing InfoFolders

The default InfoFolder available with the Enterprise Reporter software points to its sample data the "Tailor Made Shoes Company." This demonstrates how a logical, hierarchical structure can be displayed for the users.

To access InfoFolders from the Report Builder, click on **Data → Manage InfoFolders**. The InfoCenter window, as shown in **Figure 3,** will appear displaying the default InfoFolder contents.
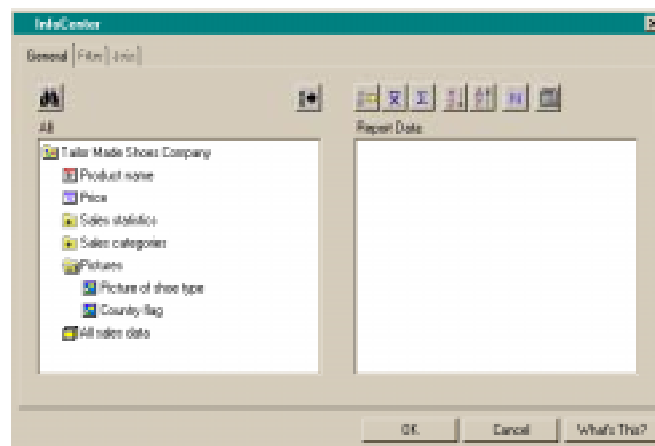

Figure 3 *InfoCenter Window*

The left-hand box, named All, in the InfoCenter contains information about the contents of the InfoFolder. Icons to the left of these items are provided to represent their types.

The various types of InfoFolder items are as follows:

- Name of InfoFolder (open/close it with a mouse click)
- Logical or physical folder (open/close it with a mouse click)
- MDDB structure (open/close it with a mouse click)
- Character data item
- Data item containing the name/location of a picture
- Numeric data item
- Date data item (numeric column with a SAS date format)

- Time data item (numeric column with a SAS time format)
- Date-Time data item (numeric column with a SAS datetime format)

Creating an InfoFolder

Now, explore how the InfoFolder can be created.

Creating an InfoFolder involves:

- *Creating a physical description of data* by assigning a name and an optional description, selecting a join strategy and selecting data from various servers.
- *Creating a logical description of data* by selecting tables & columns that the users can work with, organizing the physical data structure to a logical hierarchy, assigning attributes to the data items and specifying relations to join data tables if not done automatically.
- *Distributing the InfoFolder* by specifying which users should have access to the InfoFolder and setting limits on the amount of output users can create

To create an InfoFolder, select **Tools → Create InfoFolder** from the Report Builder. The Create InfoFolder window will appear as shown in **Figure 4**.
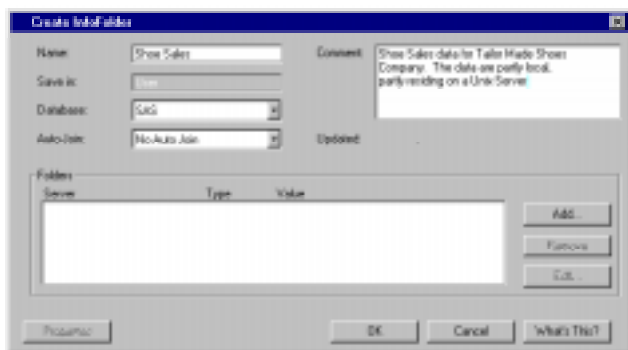

Figure 4 *Create InfoFolder window*

Fill in the following:

Name: Name of the InfoFolder. For this example, enter "Shoe Sales."

Save in: Name of folder to save in. The default is saved in "User."

Database: Type of database. In this example, a SAS View is used to point to Oracle data. Select "SAS" as the database.

Auto Join: Type of data joins for the given InfoFolder. The default is to automatically join tables within a folder. Other options to override this functionality are available.

Comments: Comments about the InfoFolder and its contents. For this example, fill in the text, "Shoe Sales for Tailor Made Shoes Company. The data are partly local, partly residing on a Unix Server."

Click the **Add** button in the Create InfoFolder window.

Pointing to the Corporate Data

In the previous section, the remote and local server locations were defined. The next step is to point to the physical data source. The Add window opens displaying the servers that are defined to the Reporter.
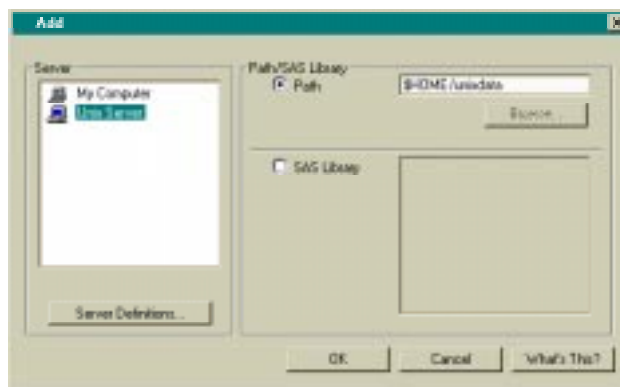

Figure 5 *Add Window that points to the corporate data or warehouse*

Select "Unix Server" in the Server window and fill in the physical path to the data in the Path/SAS Library section of the Add window. For this example, select **Path** and enter "$HOME/unixdata." An alternative is to select the SAS library by libname. This option requires that the Autoexec option be used to initially access the data source.

Select **OK** when done. The Create InfoFolder window will re-appear containing information about the new InfoFolder contents, as shown in **Figure 6**.
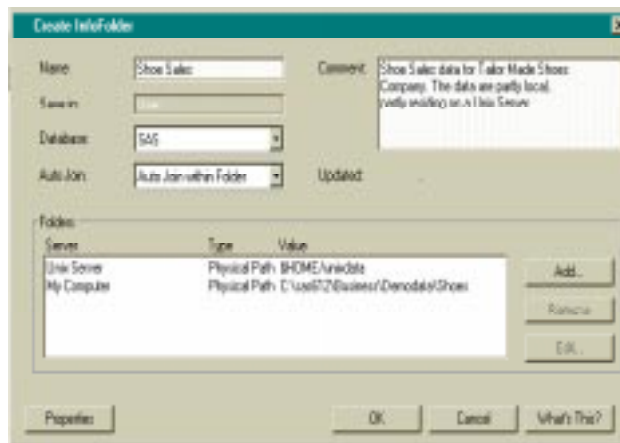

Figure 6 *Create InfoFolder window containing information about the new InfoFolder*

Repeat this process for each new path or library needed to include in the InfoFolder by clicking the **Add** button. In this example, add the local data source as shown in **Figure 6**.

Clicking on **OK** or the **Properties** button in the Create InfoFolder window will create the metadata.

Customizing Folders from the Properties window

By clicking on the Properties button, the InfoFolder can be customized. Customizations include specifying the availability of data items, how the data and items can be organized and how the tables can be joined. Click on the **Properties** button.

The physical structure of the data in the All Folders box is displayed in the General Tab of the InfoFolders Properties window. The full path names for both the local data, as well as the Unix data, will be shown.

The structure can now be reorganized and logical hierarchies made of the data. Drag items (folders, table names, MDDBs or individual columns) to the Customized Folders box. In the Customized Folders box, create new logical folders that organize items in a way best suited for the business user.

Creating a Logical folder

To create a logical folder, select the folder in the Customized Folders box from which the new folder should become a sub-folder.
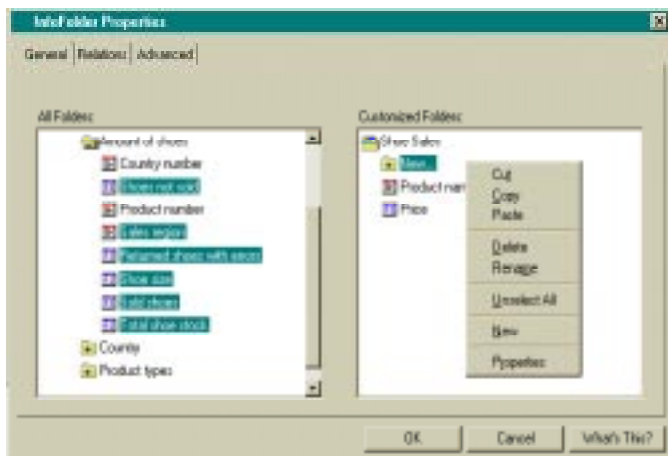

Figure 7 *InfoFolder Properties Window: Customizing the InfoFolder*

Click the right mouse button to get a pop-up menu and select **New**. A *New...* folder is presented. By selecting it and right clicking the mouse, select the properties to assign a name and other characteristics to the folder. Select **Properties** from this pop menu.

Customizing Data Items

Assign custom attributes to data items such as name, description, default statistic(s), format and sort order from the Item Properties window as shown in **Figure 8**.
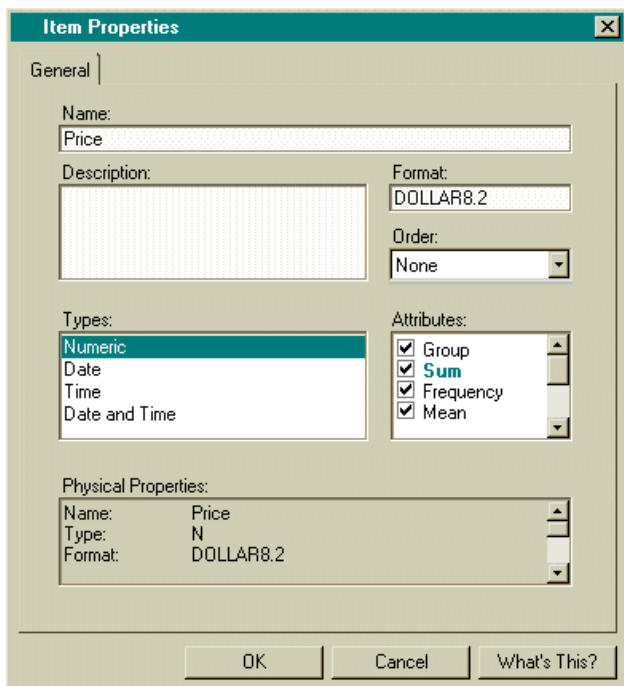

Figure 8 *Item Properties Window*

Click on **OK** to return to the InfoFolder Properties window and then click on **OK** in the InfoFolder Properties window to return to the Report Builder.

…More on the InfoFolder Properties Window

The logical organization of the reporting data can be defined in the InfoFolder Properties window (see **Figure 7**). For example, the Relations Tab provides a visual means to perform specific joins on the tables in the warehouse.

Additionally, the Properties window contains an Advanced Tab. The Advanced Tab provides an alternative method to grant specific users access and various Governor options, such as, limiting the number of input and output that is allowed for a particular InfoFolder.

Reading SAS Data Warehouse Metadata

Metadata can be exported from the SAS Warehouse Administrator Software directly into an InfoFolder in Enterprise Reporter. Export the metadata from the SAS Warehouse Administrator to a specific location. The resulting exported data contains three tables needed by the Enterprise Reporter. The three table names are: COLUMNS, GENERAL and LOCATION.

To read this data into the Enterprise Reporter, create a new InfoFolder. In the Create InfoFolder window, press the **Add** button. In the *Path* field located in the Add window, specify the location of the exported metadata by typing it in or by using the **Browse** facility. Click on the **Properties** button to create the InfoFolder. The created InfoFolder will be based on the information contained in the warehouse folder. Note, the Enterprise Reporter automatically analyzes the contents of the Data Warehouse Administrator data to create its metadata, so it is important not to mix other data sources in the same folder.

Reading any Data Warehouse Metadata into Reporting Metadata

To decrease the maintenance burden, it is recommended that the SAS Warehouse Administrator software is used to build the warehouse. The Enterprise Reporter software, however, can read any data warehouse metadata into InfoFolders automatically. A SAS software program needs to be written to convert the data warehouse metadata into a form like the Dictionary tables found in the SAS System. These tables are: TABLES, COLUMNS, and MEMBERS. Optionally, two extra tables, RTABLES and RCOLS, can be provided with information on how joins should take place. Store these tables in a location and specify this path in the *Path* field when creating an InfoFolder. In order to do this automatic read of the data warehouse metadata, it is required that it is not mixed with other data sources in the same InfoFolder. Documentation and sample programs will be provided with the software.

Building the Report

Once the business metadata has been defined for business users, they can create reports without having to worry about the physical location and column names of the data. Instead, they simply need to design the report and then assign logical data to it.

Very briefly, see below for an example of how anyone can create a report using the InfoFolder. **Figure 9** illustrates the Report Builder environment and some of its major components available to create a report. A numeric list is provided as a reference to these components.
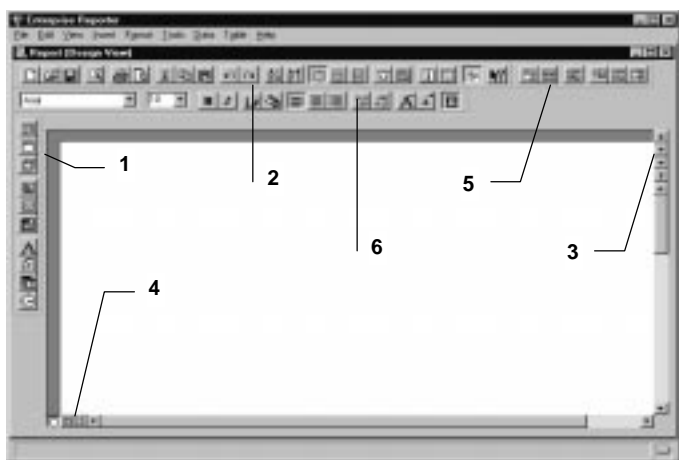
Figure 9 *Report Builder Environment*

1. Object Toolbar for inserting report objects
2. Standard Toolbar for access to other components (see Figure 11)
3. Scroll arrows to scroll through groups of data
4. Switches between Design, Data, and Print Preview
5. Table Toolbar to use on the Table report object
6. Format Toolbar for defining fonts, colors, etc.

Figure 10 shows the Standard Toolbar (Number 2 in **Figure 9**) accessible from the Enterprise Reporter Report Builder.
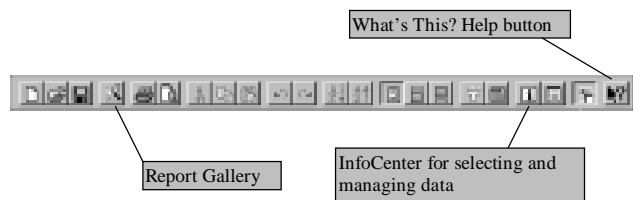

Figure 10 *Enterprise Report Builder Standard Toolbar*

To begin building or designing a report, use the object toolbar or the Report Gallery. An illustration of the example report is provided in **Figure 12.**

Then, retrieve the data source from the InfoCenter that was just created by following these instructions:

- From the Report Builder, select the **InfoFolder** icon
- From the InfoFolder selection, double click on "Shoe Sales"
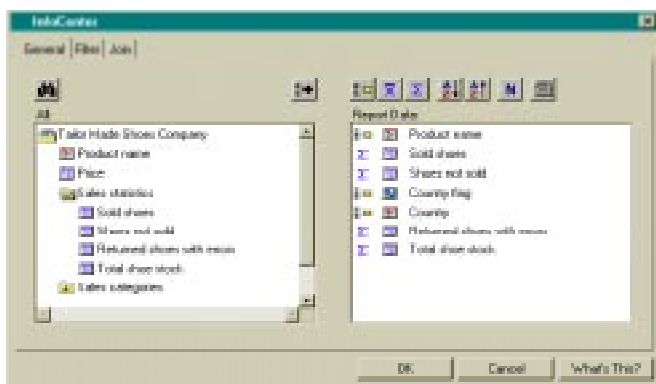- The InfoCenter appears for selection of columns to work within the report.

Figure 11 *InfoCenter window displaying the "Shoe Sales" InfoFolder*

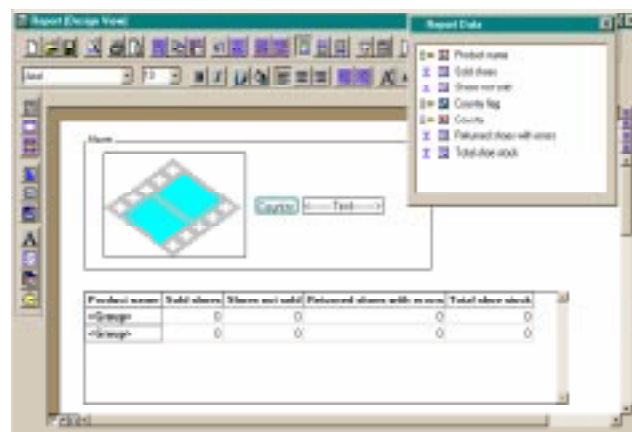A Report Data window will appear, enabling items to be dropped onto the report, as shown in **Figure 12**.

Figure 12 *Report Builder with designed report and Report Data window.*

Click on the **Print Preview** button (lower left-hand corner of the Report Builder) to see the end result. Click on the **Zoom** button to magnify the view. **Figure 13** shows the report as it will print on hard-copy.
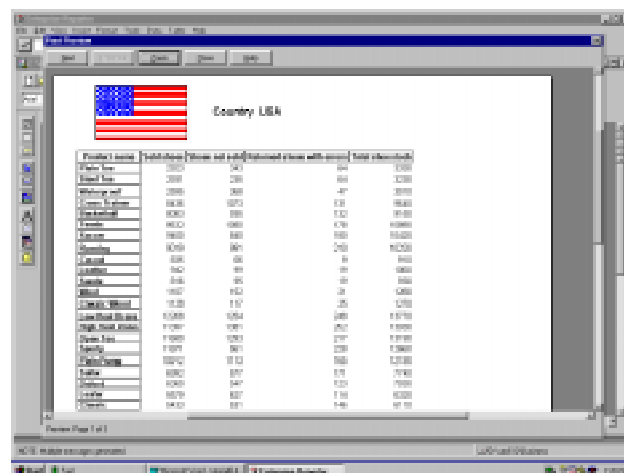

Figure 13 *Print Preview Window of newly created report*

Note, this report may be saved out in a number of different formats. Here is a list of valid saved formats:

- *Enterprise Reporter Report* (.ERP)-report definition with a pointer to the data
- *Enterprise Reporter Template* (.ERT)-report definition only
- *Enterprise Reporter Fast File* (ERF)-report with its associated data
- *HTML Document*
- *Postscript File*

Conclusion

Each day businesses collect and process large amounts of data- much of which is discarded or ignored because the organization in unable to extract meaningful information. The success of a data warehouse or access methodology is measured by the perceived usefulness of the obtained information gathered by the users. Often times the physical layout and technical metadata does not enable the business users to get the key answers they need to make daily business decisions. However, with the assignment of business metadata, the users feel comfortable with their data because it makes sense. With clearly defined business terms and business

metadata, used in tandem with a SAS data warehousing solution and the Enterprise Reporter, business users can produce the high quality reports required to make better business decisions.

References

Enterprise Reporter™ Software: User and Data Administration, Course Notes

Data Warehousing: A Blueprint For Success

A SAS Institute White Paper: Rapid Warehousing Methodology

The Enterprise Reporter™ On-line Help Facility

Special Thanks

The Data Warehousing Solution and Enterprise Reporter Teams