

ODBC: TRANSFERRING DATA BETWEEN SAS® SOFTWARE AND OTHER APPLICATIONS USING ODBC

Randall Cates, SAS Institute Inc., St. Louis, MO

Abstract:

ODBC is a powerful tool available to programmers who need to move data between different software applications. Originally pioneered by Microsoft, ODBC has become a standard method of data transfer for many software applications including the SAS software.

ODBC has three primary components. The ODBC Drivers identify a specific type and location of data. Drivers are defined through the ODBC Administrator, a program that resides and is a part of the underlying Windows operating system. The ODBC Administrator accesses the ODBC Driver Managers for information needed to create the various specific drivers. ODBC Driver Managers are specific to the different software applications.

This presentation will demonstrate how to configure different ODBC drivers, and show how to use the configured ODBC drivers to obtain SAS datasets from Microsoft Access 97, and Microsoft Access datasets with SAS PROC SQL code.

Introduction:

ODBC (Open Data Base Connectivity) is one of a number of different methods of transferring data between applications using SQL. This presentation describes what ODBC is and how it works. The presenter will then demonstrate how to:

- set up ODBC drivers for data transfer,
- transfer data from Microsoft® Access tables to SAS using PROC SQL,
- transfer data from SAS datasets to Microsoft® Access.

System Requirements (for MS-Windows):

(Other dataset to SAS)

- Microsoft Windows 3.1®, Windows for Workgroups®, Windows 95®, Windows NT®, or Windows 98®
- Base SAS Software® for the PC
- SAS/ACCESS®
- ODBC drivers for other software

(SAS to other software package)

- Microsoft Windows 95®, Windows NT®, or Windows 98®
- Access to SAS data sources
- An ODBC compliant software package

- The SAS Universal ODBC driver

What Is ODBC?

ODBC is a programming interface which enables applications to access data from database management systems that use Structured Query Language (SQL) as a data access standard. ODBC was pioneered by Microsoft® for the Windows environment and has been accepted by many software companies including Oracle®, Ingres®, Lotus®, Powersoft®, and SAS® Institute Inc. The tremendous acceptance of this standard has also prompted its adoption on other platforms including Macintosh®, OS/2®, and Unix®.

ODBC, therefore, offers the great advantage of connecting directly to data from many previously unavailable sources, even from different networks. And the same set of code can be used to access most of the different databases.

There are three components to ODBC existing on your operating environment as illustrated in Figure 1: the ODBC Administrator, the different ODBC Driver Managers, and the different ODBC Drivers. Notice that the ODBC Administrator sits aside from the direct line between your SAS program and the data.

How Does ODBC Work?

To use a comparative analogy, the old methods of transferring data (re-typing, flat-files, etc.) are like a truck and trailer and ODBC is like a door and a key.

The truck and trailer approach required that you enter the application that was sending the data and either download the data to a flat-file, or print it out (like loading a truck). Next you had to enter the recipient application and either create code to read in the flat-file, or re-enter the data from the printout (like unloading the truck). These processes were inefficient in three ways. First, they required significant amounts of time and typing. Second, they required either, one programmer having access to and knowledge of both applications, or the skills of two programmers. And third, there was a much greater potential for data errors.

More third-party applications have become available that do part or all of the process for you. But this required the purchasing of yet another software package.

With ODBC you need access to the ODBC Administrator which is integral with all current Microsoft Windows Operating Systems, and an ODBC Driver Manager and ODBC Drivers for the data source application. You don't actually need (in most cases) the originating application to obtain access to the data.

First, you open up the ODBC Administrator and, either create a new ODBC driver, or modify an existing one to obtain access to the data. It's like making a key!

Then in your receiving application, you import the data using the ODBC driver, which functions as a code translator to get your request for data to the database.

One final note. ODBC comes in 2 distinct flavors: 16-bit and 32-bit. Windows 3.x and applications designed for Windows 3.x can only use 16-bit ODBC drivers. Windows 95, Windows NT and newer applications primarily use 32-bit ODBC drivers.

Setting Up Drivers:

ODBC Drivers for other applications

Most compliant applications include ODBC driver software with their installation package. If you don't have the particular application, or the application did not include the driver software, you need to contact the software company to obtain it. Most software companies have ODBC driver installation software available for free download off their web sites. It is also recommended that you check their web site for the latest versions.

SAS ODBC Drivers

SAS Institute Inc. provides two possible ODBC drivers.

The installation disk for Base SAS Software for Windows comes with a floppy disk containing the SAS ODBC driver software. Alternatively you may download an updated version from the SAS web site. This driver will allow another application in most cases to read, write, or modify SAS datasets. However, this driver will only work on a system with Base SAS Software and SAS/Access products installed. Also, when accessing local network SAS datasets with this particular driver, a separate SAS session is started and you need to manually close the SAS session.

A second option is to purchase the SAS Universal ODBC Driver from SAS Institute Inc. This driver can be installed and used on systems without Base SAS Software and can be used to access SAS datasets both on the local system and on networks. The SAS Universal ODBC driver only allows read access to the data and is easier to configure.

The example of SAS data transfer in this presentation uses the SAS Universal ODBC Driver.

Configuring the ODBC Drivers for Use:

In order to access external data using ODBC drivers, you must first configure a driver to point to the data and define a method of access. This is done through the ODBC Administrator.

You can find the icons for accessing the ODBC Administrator in a number of places on your Windows screen. However the most consistent place would be: from the Start button, go to the Settings menu item and select Control Panel. Find and click on the ODBC icon. This brings up the ODBC Data Source Administrator (Fig. 2). This application interface will vary depending on your system, but all will have a window containing any currently defined ODBC drivers, including some generic drivers available for modification, and Add, Remove, and Configure buttons.

Click on the Add button, which will bring up the Create New Data Source dialog box (Fig. 3) where you can choose from the currently available ODBC drivers. If the driver that you want is not listed then you will need to install it.

Select the type of driver that you want from the list and click on the Finish button (circled). In this example the SAS Universal ODBC Driver item is selected. You are then brought into the software's Driver Setup dialog box (Fig. 4)

In the Data Source Name field enter some name that you will use to refer to when calling for the data. Here we've used the name "MWSUG ODBC Example." The Description field is filled in for you, or you may use this for more pertinent information. You then need to fill in the SAS File field with the correct path. You may wish to access only one dataset by filling in the entire path name including the dataset name, or access a directory of datasets by typing an asterisk for the dataset name, as in our example.

Finally, there is an Options button that expands the dialog box to include fields for User IDs and Passwords (not shown).

Once you've entered the necessary information, click on the OK button, which brings you back to the Data Source Administrator. The new, configured ODBC

Driver is added to the list. Finally, click on the OK button to save the changes.

Configuring other data sources for access by SAS is almost the same. The individual software ODBC driver setup dialog windows will differ. Figure 5 shows the ODBC Driver setup window for Microsoft Access 97®. Again, there are fields for Data Source Name and Description. The Database field has been replaced by a box with a series of buttons to access the path names to the data. The Data Source Name has been entered as "Northwind" and the Database has been set to point to the Microsoft sample Northwinds database.

Transferring MS-ACCESS Tables to SAS:

Once you've defined ODBC drivers to access your data, it's a relatively simple matter to read the data into SAS through PROC SQL statements. Using the above MS-Access example, the following code first sets a connection to the Northwinds database. Then the code creates the SAS dataset xxxx.products from the Northwinds table titled products.

```
LIBNAME XXXX "C:\My Documents";
```

```
PROC SQL;
CONNECT to ODBC as North
      (DSN = Northwind);
CREATE TABLE xxxx.products as
      SELECT * from connection to North
      (select * from Products);
QUIT;
```

The above code results in a SAS dataset with 77 observations and 10 variables, which correspond to the rows and columns of the Northwinds table that was accessed.

With ODBC and PROC SQL code, it is possible to limit the amount of data accessed. The following code limits the number of rows based on the number of Units in Stock.

```
PROC SQL;
CONNECT to ODBC as North
      (DSN=Northwind);
CREATE TABLE xxxx.lowprods as
      SELECT * from connection to North
      (SELECT * from Products where
      UnitsInStock < 30);
QUIT;
```

Another possibility available is the ability to join related tables together into one resultant dataset. The following code merges the Customers table with the Orders table by the CustomerID key variable.

```
PROC SQL;
CONNECT to ODBC as North
```

```
      (DSN=Northwind);
CREATE TABLE xxxx.custord as
      SELECT * from connection to North
      (SELECT * from Customers INNER
      JOIN Orders
      ON Customers.CustomerID =
      Orders.CustomerID);
```

```
QUIT;
```

Finally, data transfer using ODBC may also be done through the SAS/ACCESS Query windows. It is not the purpose of this presentation to show this method however.

Transferring SAS Datasets to MS-ACCESS:

When transferring SAS data on your local machine to other applications using the SAS Universal ODBC Driver, it's not necessary to have a copy of Base SAS Software. You can install the ODBC driver software, configure a driver for your data, and then import the data to the other application.

Again, using the prior example driver configuration (Fig. 4), we can now open MS-Access and start adding the data. From the File menu, select the "Get External Data" menu item, then the "Import" submenu item. This will bring up the Import dialog box (Fig. 6).

To access ODBC driver data, move to the bottom left-hand field titled "Files of Type:" and select the "ODBC Databases" entry at the bottom of the list. This brings up the Select Data Source dialog box (Fig. 7). Select the configured driver and click the ok button. There are two file tabs at the top corresponding to the two different paths discussed earlier, i.e. a directory or a dataset.

If the driver was configured to access a directory rather than one particular dataset, a new dialog box appears with the available datasets in that directory. Choose one and a new table is created in MS-Access from the SAS dataset.

If the driver was configured to access one particular dataset, then it will not be listed on the above dialog box. Click on the Machine Data Source file tab at the top of the dialog box to obtain the list of drivers accessing individual datasets. From this list choose the appropriate dataset and the corresponding table will be created.

Conclusion:

ODBC is a powerful tool available to programmers who need to move data between different software applications. Originally pioneered by Microsoft, ODBC has become a standard method of data transfer for many software applications including the SAS software. It is available on four different

operating systems; MICROSOFT WINDOWS (most flavors), OS/2, MACINTOSH, and UNIX. It has a number of advantages over the old truck and trailer methods of data transfer.

ODBC provides a common grammar (SQL) for accessing a wide variety of data types. Transfers of data can occur much more quickly and with greatly reduced chances for errors. ODBC can be used to obtain data residing on both local systems and remote servers. ODBC makes programmers' jobs more effective by speeding up data transfer and reduces their need to purchase and learn new applications.

SAS programmers can use ODBC in both directions, since SAS provides both data recipient (Client) and data source (driver) support for ODBC. Client support is provided through both PROC SQL code and SAS/ACCESS to ODBC. Driver support is provided with either the SAS ODBC Driver which needs the SAS Software to work, or the SAS Universal ODBC Driver which does not need the SAS Software installed.

ODBC works with three components. ODBC Drivers identify a specific type and location of data and are the keys that recipient applications call to obtain data from the data sources. Drivers are defined through the ODBC Administrator, a program that resides and is a part of the underlying Windows operating system. The ODBC Administrator accesses the ODBC Driver Managers for information needed to create the various specific drivers. ODBC Driver Managers are specific to the different software applications.

This presentation has demonstrated how to configure different ODBC drivers, and how to use the configured ODBC drivers to access SAS datasets from Microsoft Access 97, and Microsoft Access datasets with SAS PROC SQL code.

References:

SAS Institute Inc. (1995), *SAS ODBC Driver Technical Report: Users Guide and Programmers Reference*, Release 6.11, Cary, NC: SAS Institute Inc.
SAS Institute Inc. (1993), *SAS Companion for the Microsoft Windows Environment Version 6, First Edition*, Cary, NC: SAS Institute Inc.

SAS, SAS/ACCESS software

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

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

Acknowledgements:

The following people contributed to the preparation of this paper:

Carol A. Prombo, PhD
Judi Stotlemeyer, SAS Institute Inc.
Scott Hayden, SAS Institute Inc.

The author can be reached at:

Randall Cates
Technical Training Specialist
Professional Services Division
Southwest Region
SAS Institute Inc.
100 South Fourth Street
St. Louis, MO 63102
(314) 421-6364
racate@wnt.sas.com

Figure Appendix:

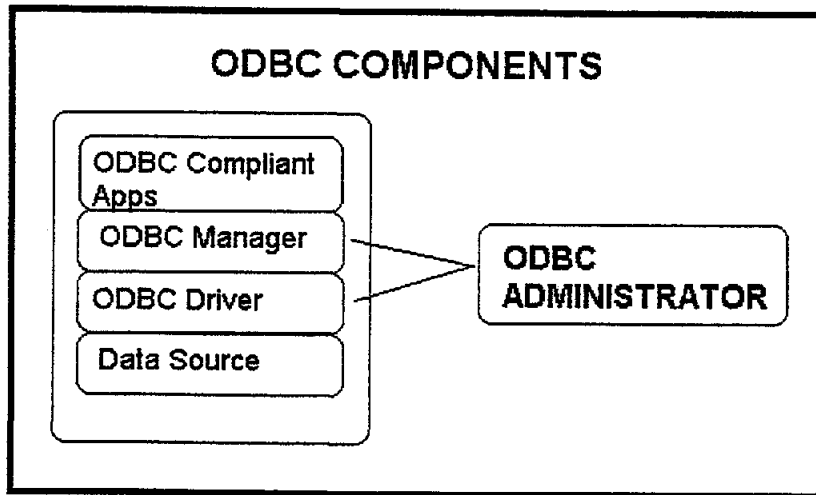


Figure 1: ODBC Components.

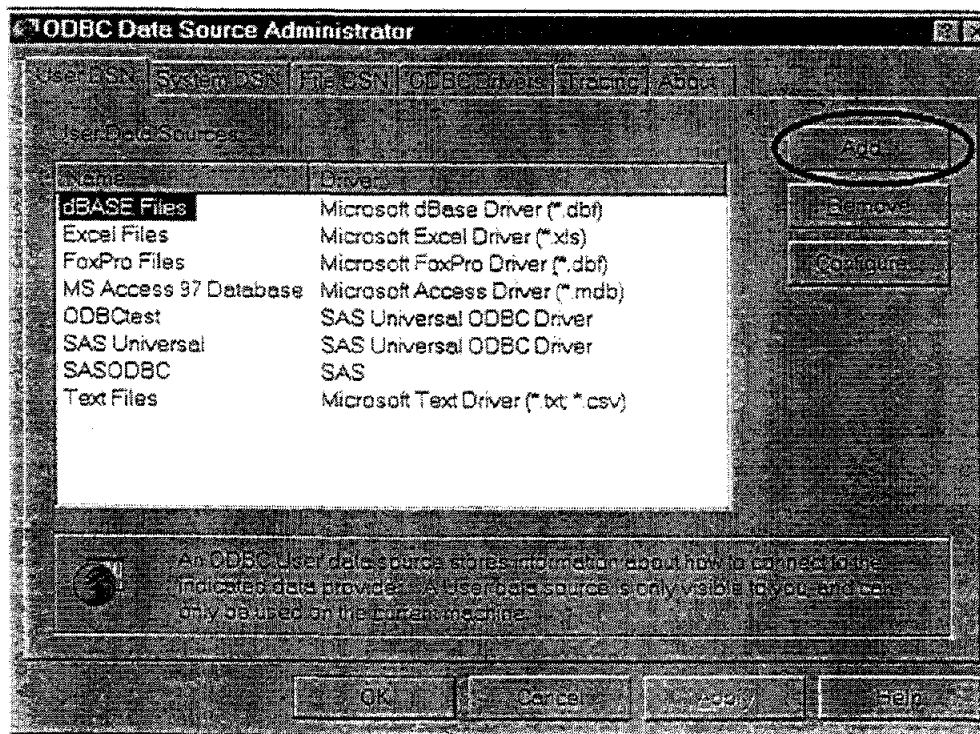


Figure 2: Data Source Administrator

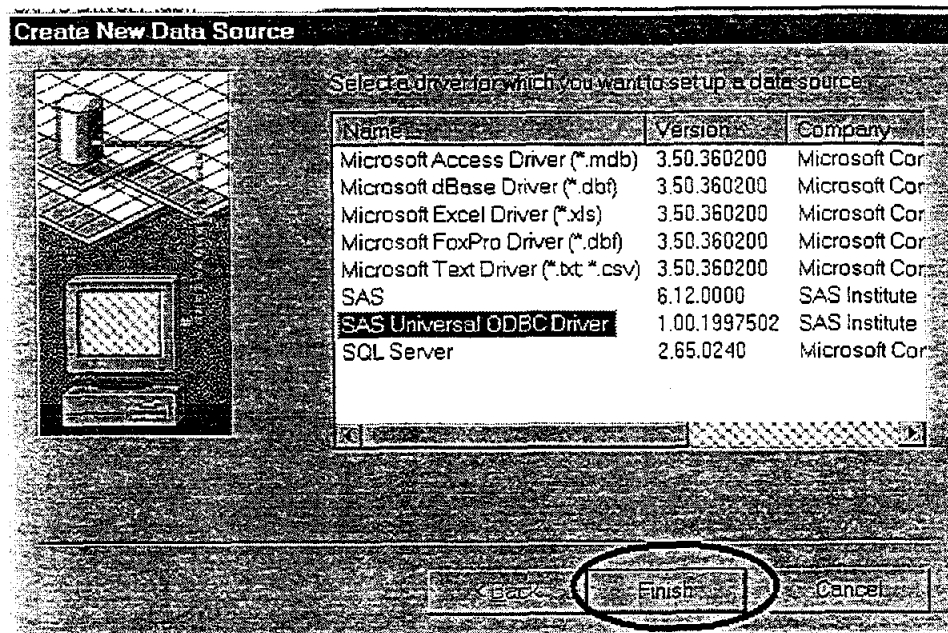


Figure 3: Create New Data Source

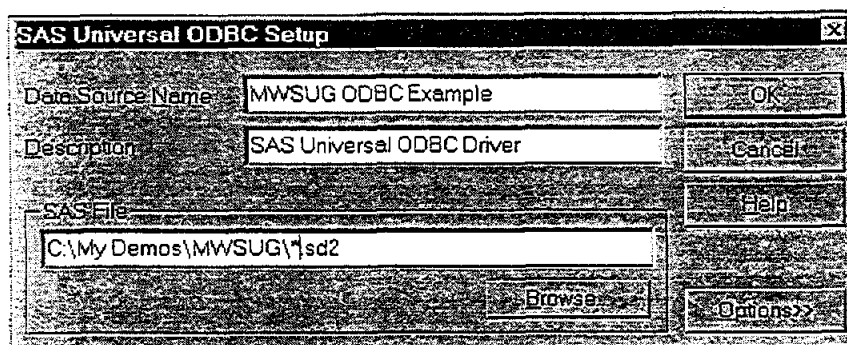


Figure 4: SAS Universal ODBC Setup

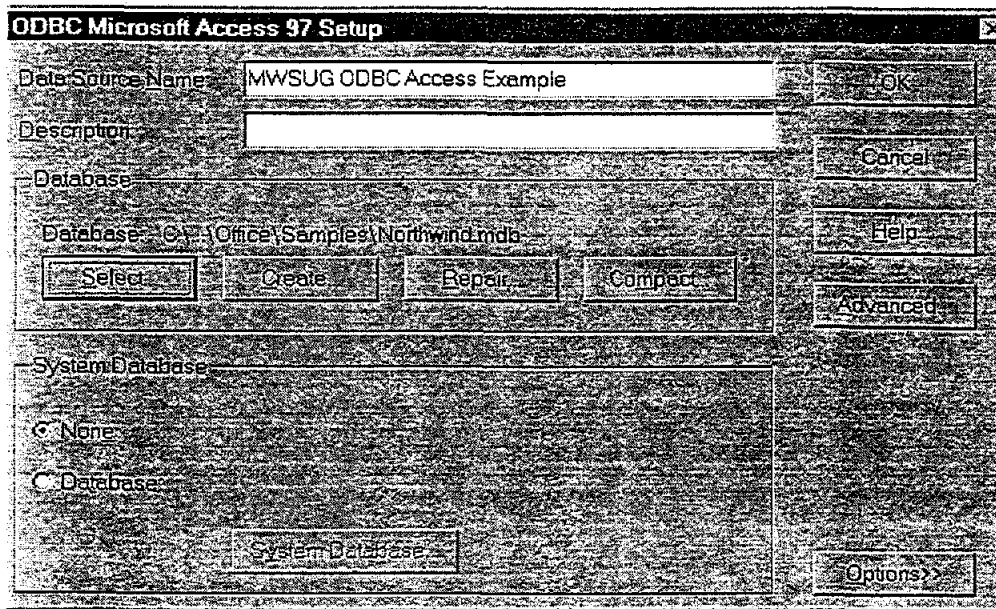


Figure 5: ODBC Microsoft Access 97 Driver Setup Window.

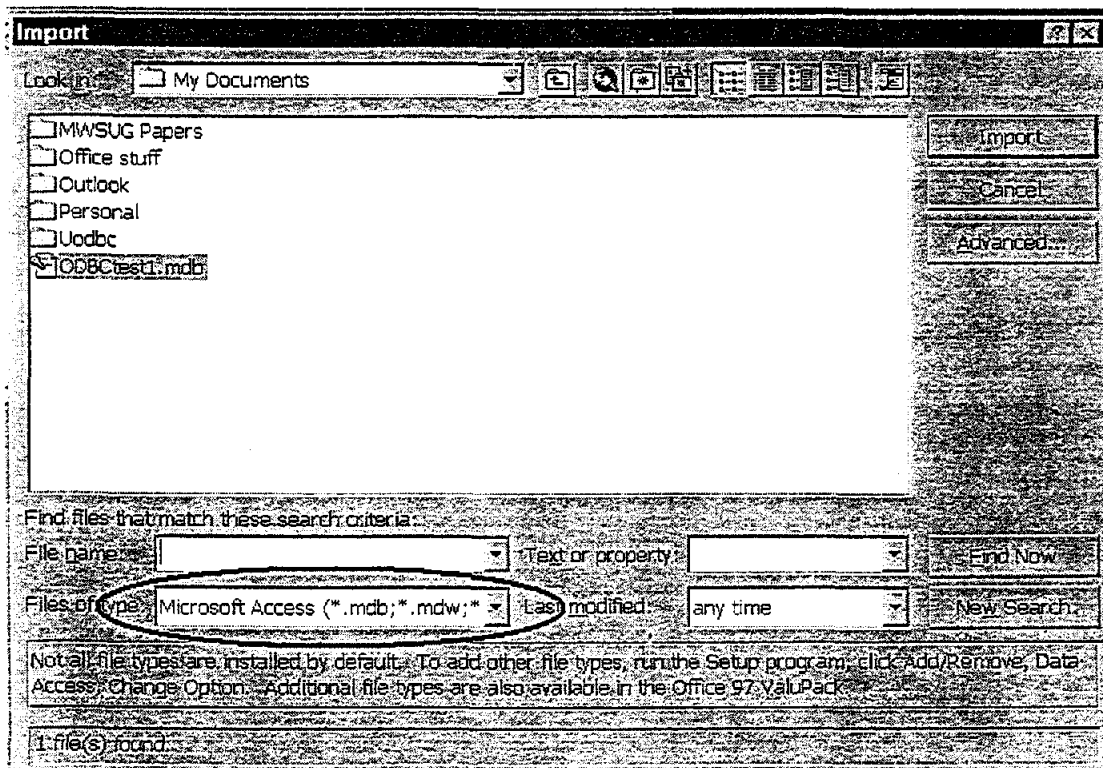


Figure 6; MS Access Import Dialog Box

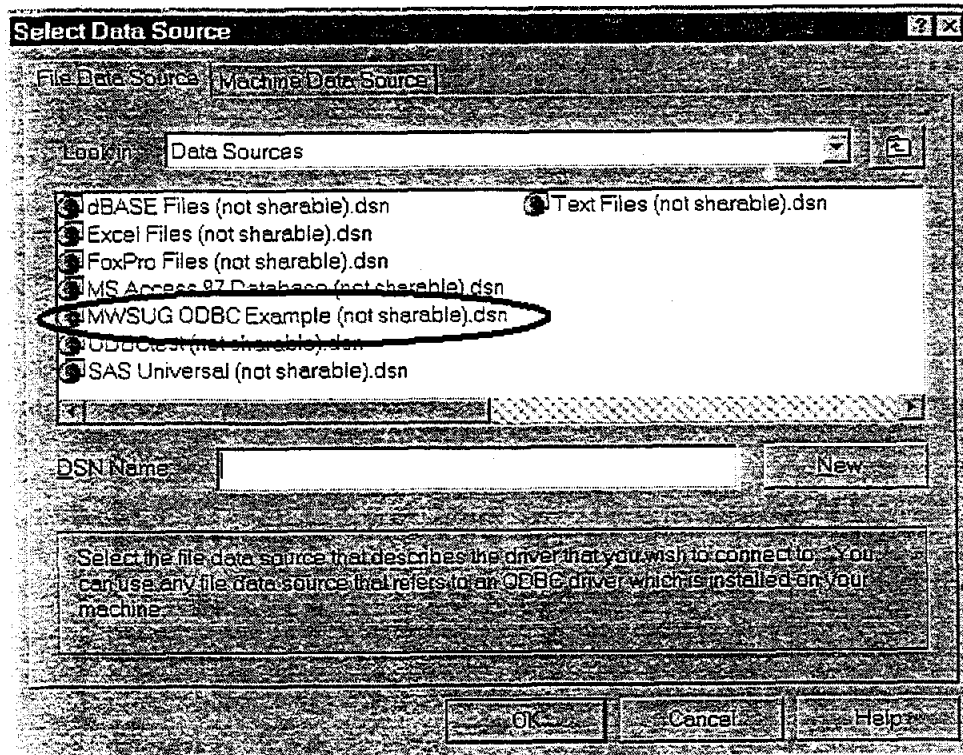


Figure 7: Select Data Source