

Paper FF-014

Tips for Moving to SAS® Enterprise Guide® on Unix Patricia Hettinger, Consultant, Oak Brook, IL

ABSTRACT

Many companies are moving to SAS® Enterprise Guide®, often with just a Unix server. A surprising number of SAS analysts are intimidated by this, even those who have used PC/SAS in the past. This paper gives some tips and tricks when moving your code into a SAS® Enterprise Guide® project, whether from PC/SAS or batch Unix. It covers the difference between outside and embedded code, transcoding differences between the Windows and Unix environments and common installation problems.

INTRODUCTION

SAS® Enterprise Guide® (EG) is a very powerful data visualization tool. One trend the author has noticed is dropping PC/SAS in favor of SAS® EG with a Unix server. This can be quite a change for users unused to a GUI interface (most users running SAS code from the Unix prompt) and PC/SAS users (what is this Unix server anyway?). It helps to understand that even though it looks interactive, you are running your code in batch and just directing your output back to your screen.

This paper gives some tips and tricks for setting your options, moving your existing code into a project, visualizing data and getting useful output out to those who need it. Also addressed will be some common installation issues. The focus is mainly moving from PC/SAS, whose users, in the author's experience, seem to be the most intimidated by SAS® Enterprise Guide®.

SETTING YOUR OPTIONS

Before you can use SAS® EG, you must set up some of your options, especially the repository and servers. You set them by dropping down the 'Tools' toolbar and selecting 'Options..':

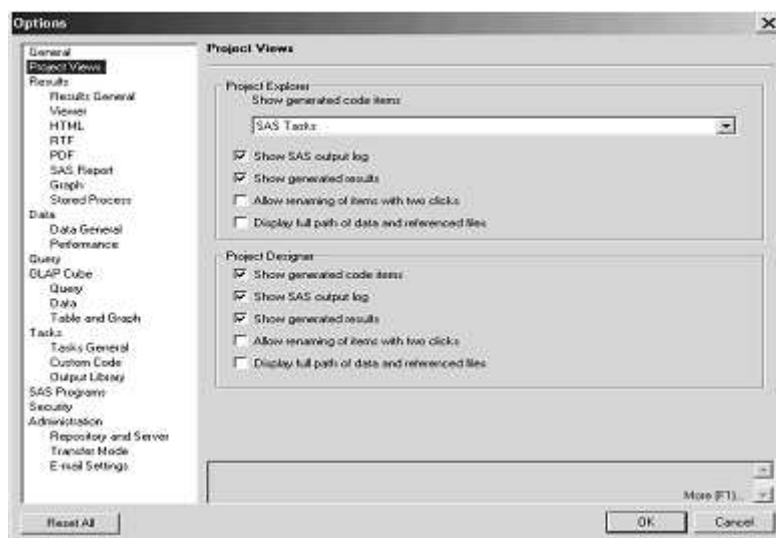


Figure 1: Options Dialog Box

You will need the location of your Unix repository and server to use SAS® Enterprise Guide®. You will also need the port number. You can select 'automatically add local SAS server to server list if available' but don't be surprised if there isn't one. Some companies are only using SAS® EG with a Unix server. For the purpose of this paper, we will call it SASMain.

Under Data>Data General, deselect 'Automatically open data when added to project'. SAS datasets on UNIX can be very large and will lock up your project if you attempt to open them. This is doubly true for Teradata or Oracle tables assigned with a libname. In the Data>Performance section, you might want to increase the number of distinct parameter values to 100.

Under Tasks > Output Library, move WORK to the top of the list as you may find you don't have write authority over the others (most likely EGTASK and SASUSERS). This will result in an error when running some of the built-in tasks.

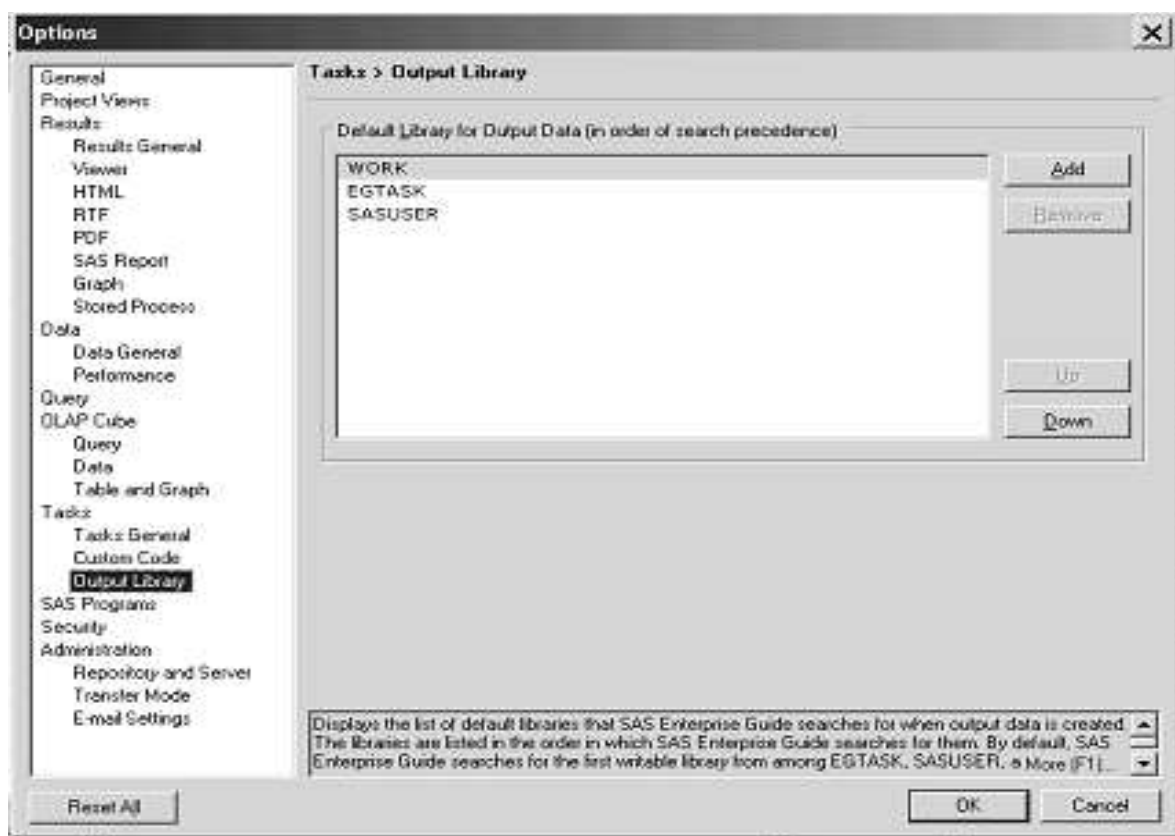


Figure 2: Moving work libraries to the top of the list

CREATING A PROJECT.

Once your servers are set up, you can create a project. When you open up SAS® EG, You should find yourself in one of two views, Project Explorer or Project Designer. Project Explorer is the only view where you can set security or turn on the project log. You do this by highlighting the project and selecting properties from the 'File' toolbar:

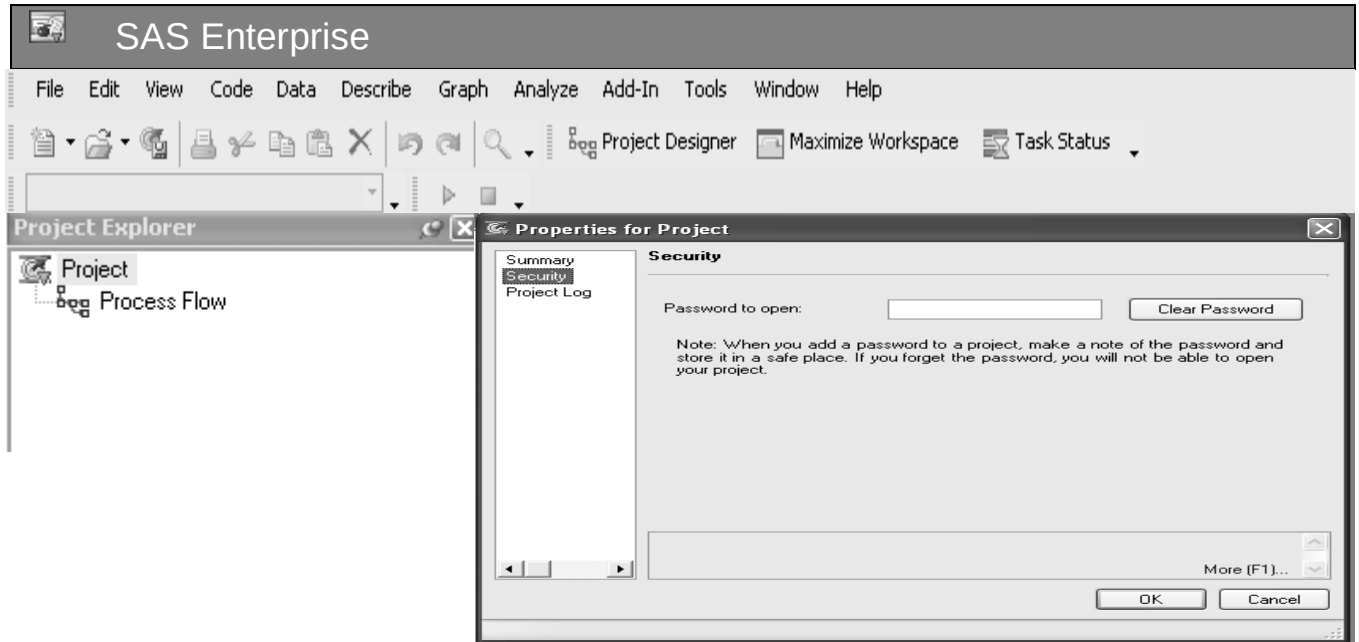


Figure 3: Project properties in Project Explorer

You will probably find it easier to work in the Project Designer:

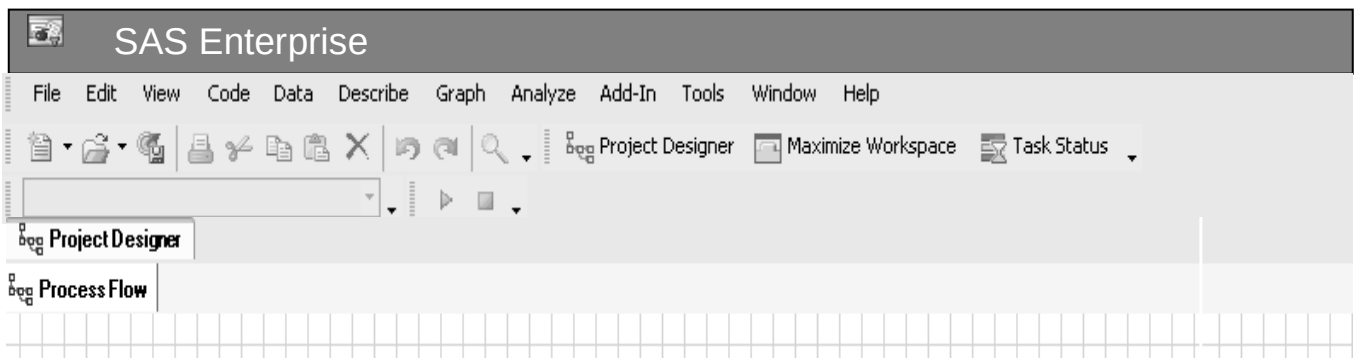


Figure 4: Project Designer view

Now suppose you have a PC/SAS program in directory C:\SAS Programs named 'Read in File' that reads in a Unix file residing on the same box as your new SAS® EG server. And let's suppose it uses a remote submit like this:

```
rsubmit;  
filename acctfile "/myfiledir/acctfile.txt";  
libname savedata "/mysasdir";
```

```
Data savedata.custinfo;  
  infile acctfile;
```

Input

```
@1 org_code          $3.  
@4 acct_num          $16.  
@20 cust_name         $40.  
@60 street_addr1      $40.  
@100 street_addr2     $40.  
@140 city             $40.  
@180 state_abbrev     $2.  
@182 zip5             $5.  
@187 zip4             $4.  
  
;  
run;  
endrsubmit;
```

To run the above on your Unix server, you would open the code within the project from your local server:

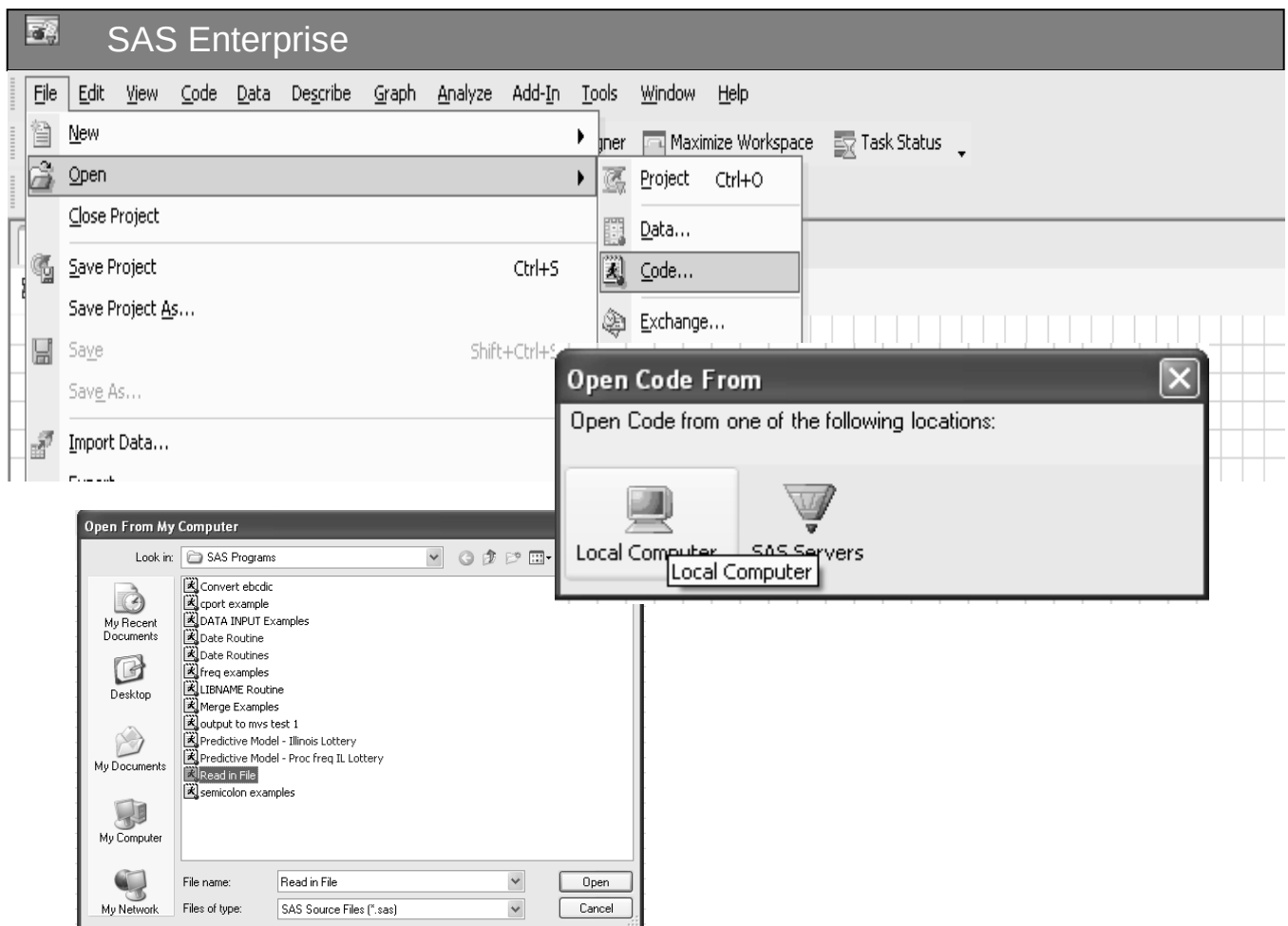


Figure 5: Opening your existing PC/SAS code

Choosing this file will add it to your project as an icon and open it up in the editor:

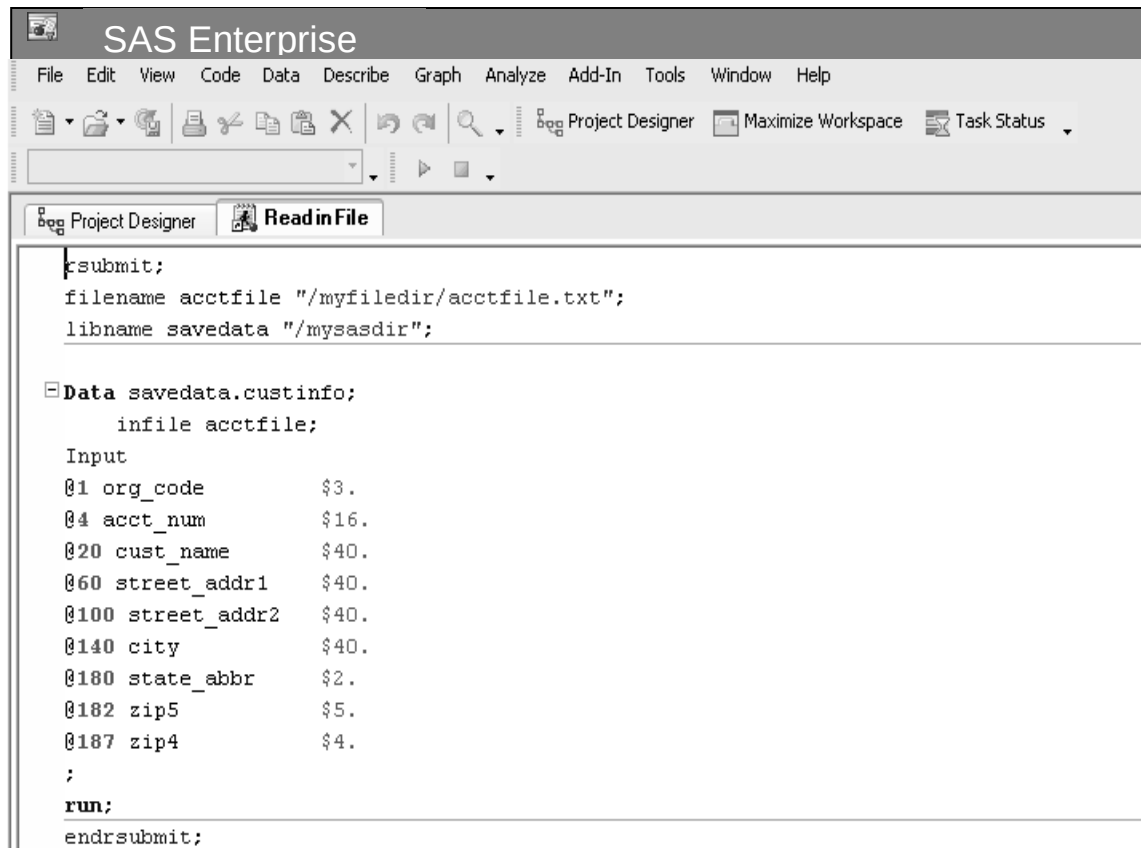


Figure 6: Editor view

To run this code on your SAS UNIX server, there are two simple changes. The first is to remove the rsubmit and endrsubmit statements in the code. The other is to select your Unix SAS server. You may select the server from the 'Code' drop down box in the editor:

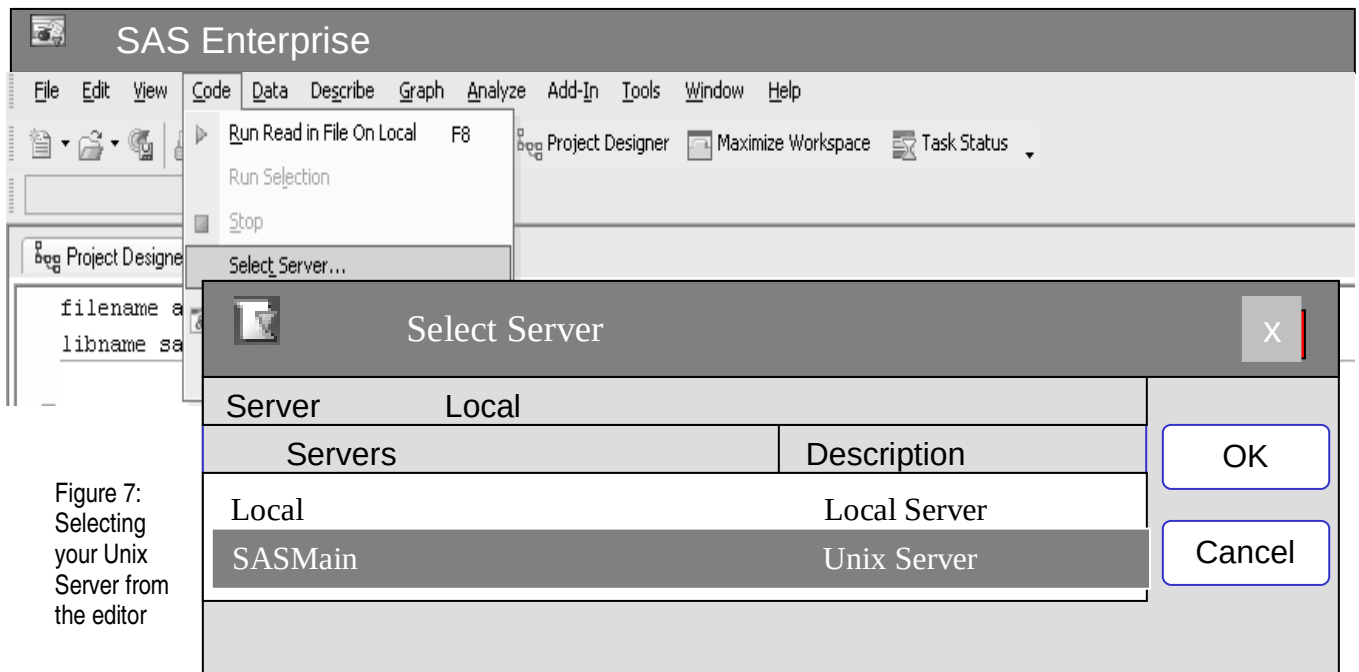


Figure 7:
Selecting
your Unix
Server from
the editor

Or by right-clicking on the code icon in Project Designer:

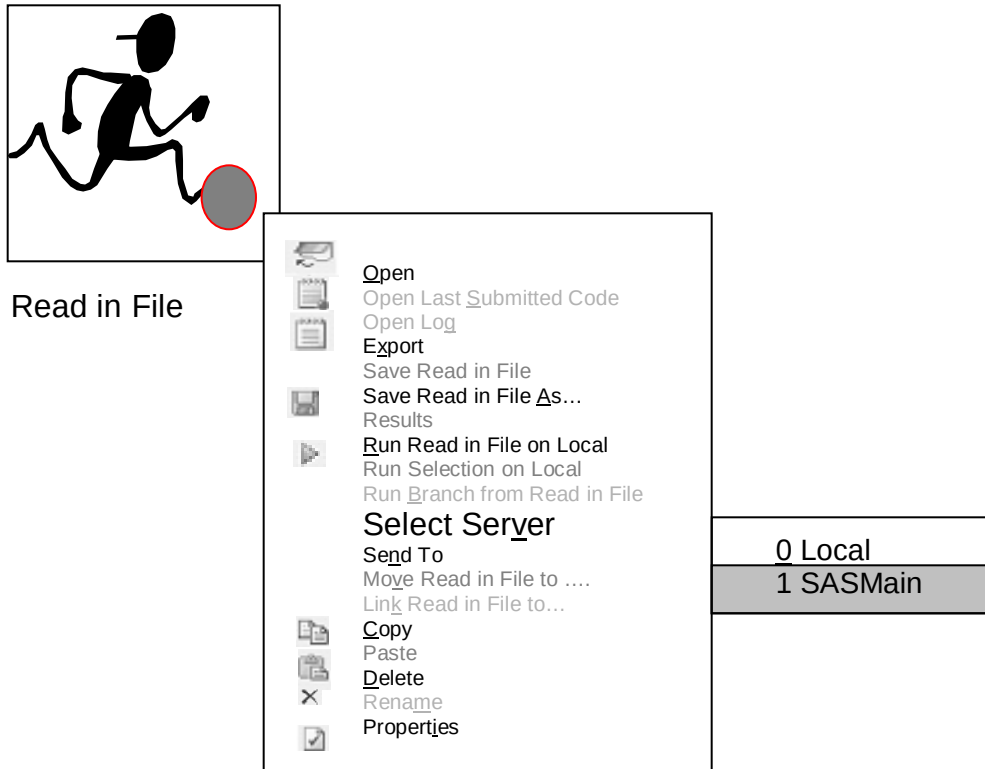


Figure 8: Selecting your Unix server from the code icon

EMBEDDED CODE VS. EXTERNAL CODE

Now that you have your code brought into the program, there are three places we could keep it. One is to leave it where it is, on your local drive. Another is to embed it in the project. A third option is to keep it in a Unix directory.

Embedded code stays with the project – changes are automatically saved when the project is saved. Embedded code is the default for new code. If the code is stored in an external file, you will be prompted to save changes when exiting either the editor or the project. You can have many copies of the same code with the same name in a project. However, only one copy can be external. The others are embedded. In this case, changes to the external code are not carried over to the embedded copies and vice versa. Changes to the embedded copies are not even carried over to other copies as illustrated below:

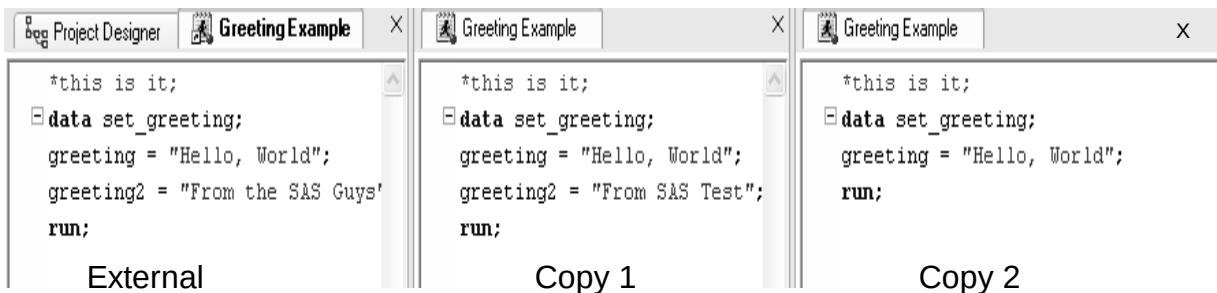


Figure 9: Multiple versions of the same program

Store your code externally to be safe. It's always possible to delete a program icon by mistake and thus lose a crucial program. If the program is stored externally, you only delete the link. External code also makes it easier to bring into other projects or even to run outside of SAS® EG.

In the example we have been using, the code is already stored on the local computer. How might you save it to a directory on the Unix server? Use the 'save as' option from either the 'File' toolbar or by right-clicking the code icon and setting the 'General' properties. You will again be prompted to save the code to either the local computer or the SAS server. Save it to the SAS server as many companies discourage storing any code or data on local drives. A drop down box very like the illustration below should appear. Click 'Save' and your code should both be stored on your Unix box and the link set up in the project.

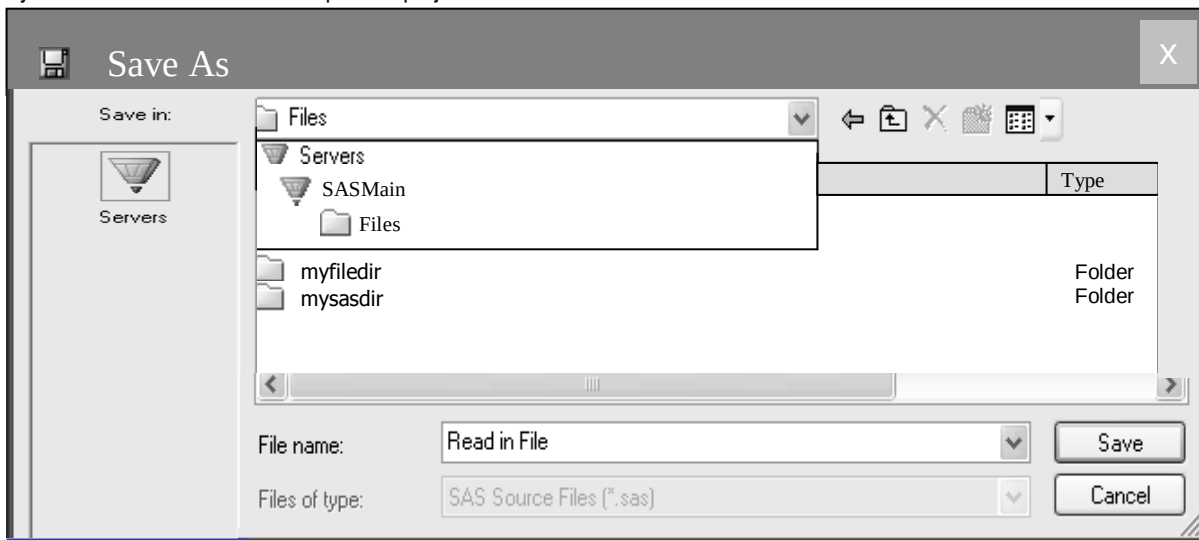


Figure 10: Saving your code to a SAS Unix server

To keep this code link in the project, you will need to save the project, preferably to your Unix server. You will see the same prompts as in the save code example. The file extension for projects is .egp.

WHAT HAPPENS IN THE PROJECT STAYS IN THE PROJECT (MOSTLY)

Program logs and report output stays in the project provided that you save the project. If not, the logs and reports will be lost unless they are exported or sent to Word or Excel. Work files will remain as long as you are in SAS® EG even if you switch projects. Once you exit SAS® EG, work files are released, just like any other version of SAS. Permanent files saved to a SAS library must be allocated each time you enter SAS, again just like any other version. There is an assign library task in SAS® EG but you may find it easier to just write the code as we did in the 'Read in File' program.

RUNNING A CODE SEQUENCE

Suppose you have now added three programs to your project and you'd like to run them in order:

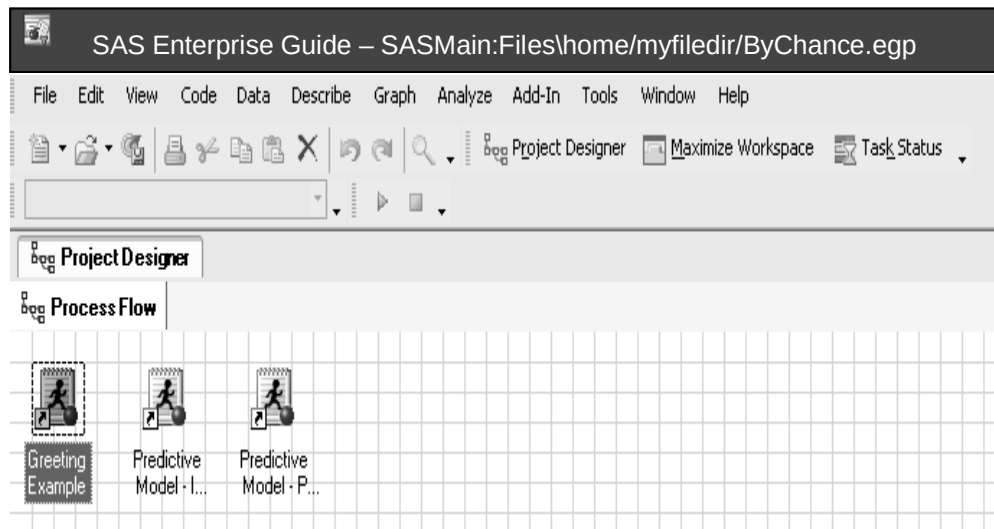


Figure 11: Three unconnected programs

You could run each one at a time by selecting the appropriate icon and either hitting the PF8 key, selecting 'Run' from the "Code" toolbar or right-clicking on the icon and selecting 'run' there. You could put some "%include" statements in the first program, 'Greeting Example' with the location of the other two programs. Or you could simply link the programs together by right clicking on the icons and linking to the desired program:

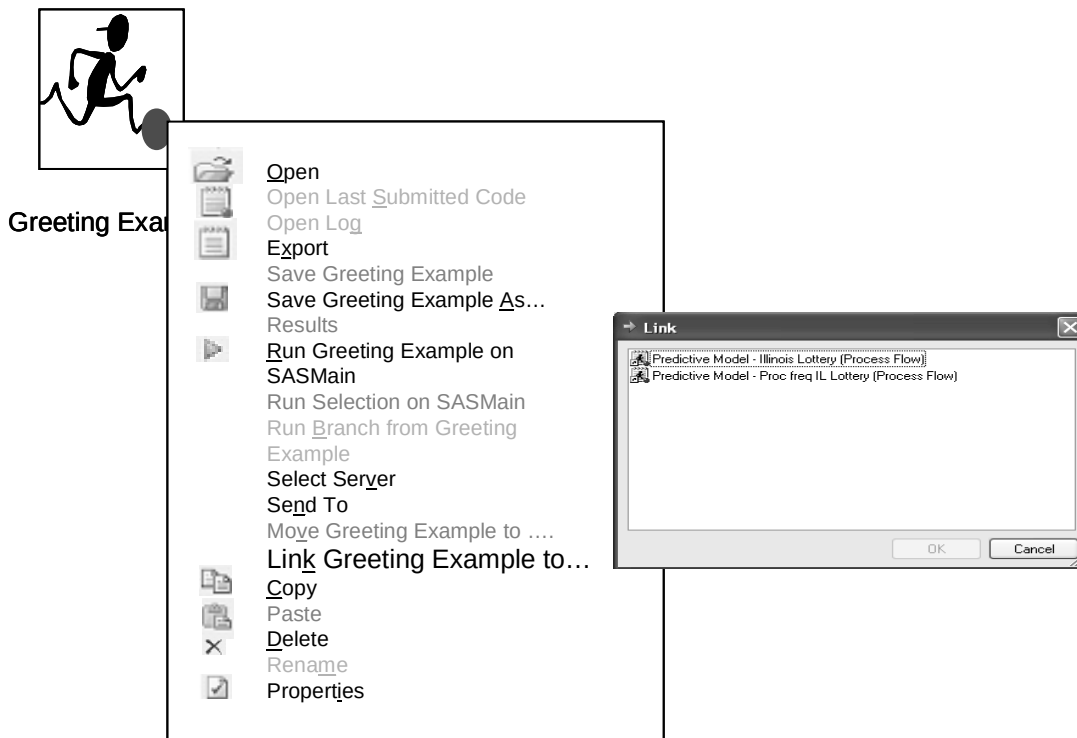


Figure 12: Connecting programs

After linking the programs together, the sequence may be run from the project, the process flow, the branch or the individual program. Note that icons highlighted in green are programs actually running, while those in yellow are waiting to run:

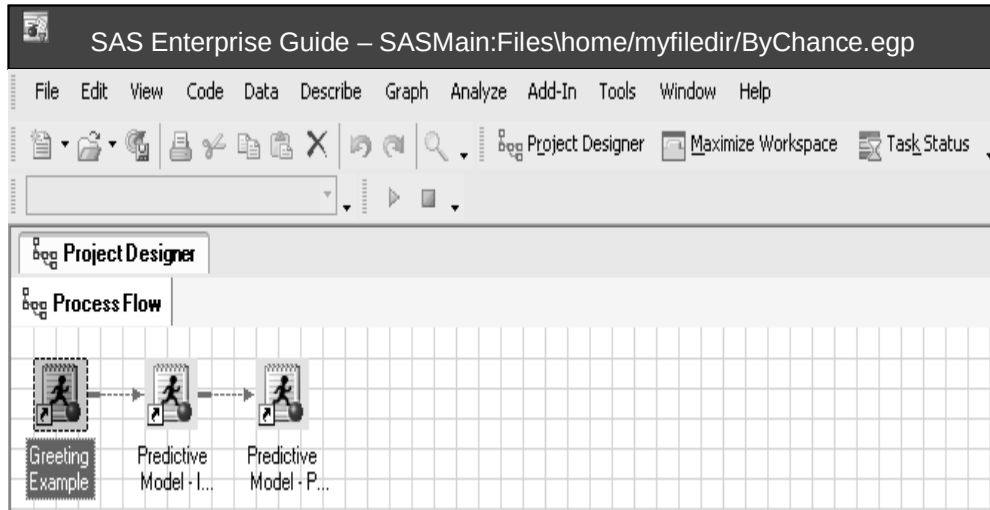


Figure 13: Project Designer view of the code while running

Task Status			
Task	Status	Queue	Server
Greeting Example	Running...		SASMain
Predictive Model – Illinois Lottery	Waiting on [Greeting Example]...		SASMain
Predictive Model – Proc Freq IL Lottery	Waiting on [Predictive Model – ..]		SASMain

Figure 14: Another view from the Task Status box

Links to the output from these programs are automatically created during execution. In this case, there are work SAS datasets and an HTML report.

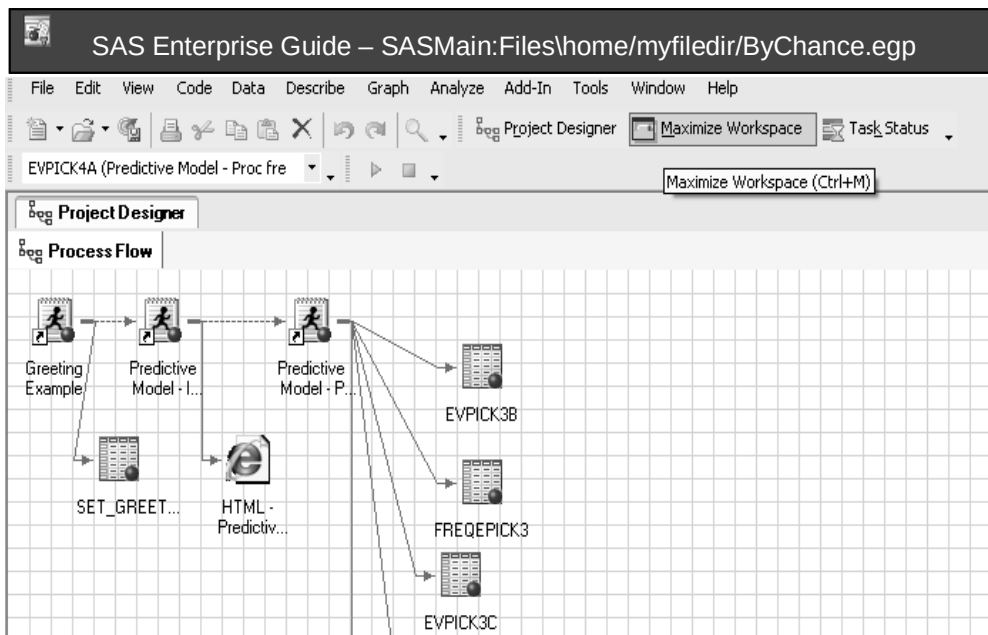


Figure 15: Output from run in Project Designer

IMPORTING SPREADSHEETS, ACCESS TABLES AND OTHER NON-SAS DATA FROM YOUR LOCAL DRIVE.

This is quite easy to do providing you don't have any characters in the data that won't transcribe to Unix. Let's assume you want to import a spreadsheet. First you would select 'Import Data...' from the 'File' toolbar. After selecting the spreadsheet from your local drive, you would be prompted for the worksheet:

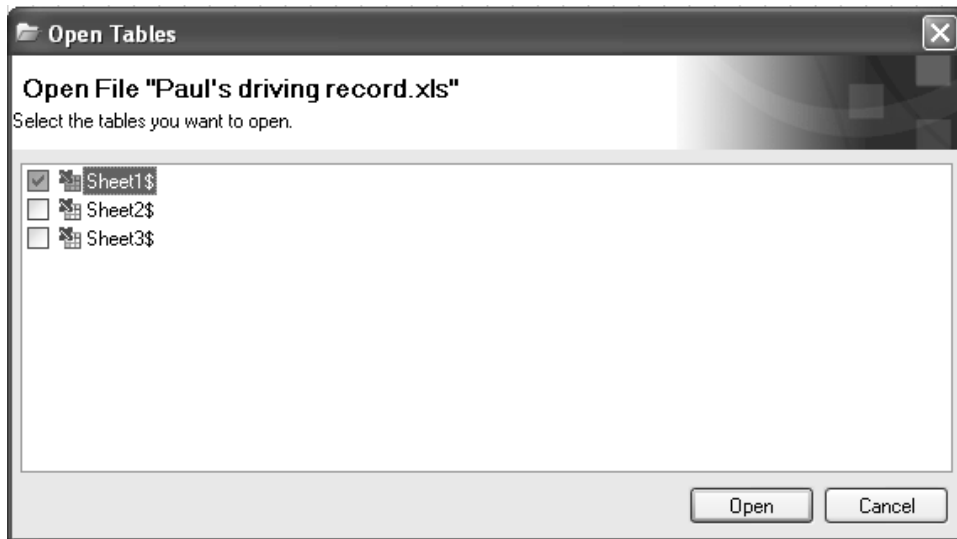


Figure 16: Initial Excel import dialog box

You then have the option of either viewing the file as it is or opening it as a SAS data set.

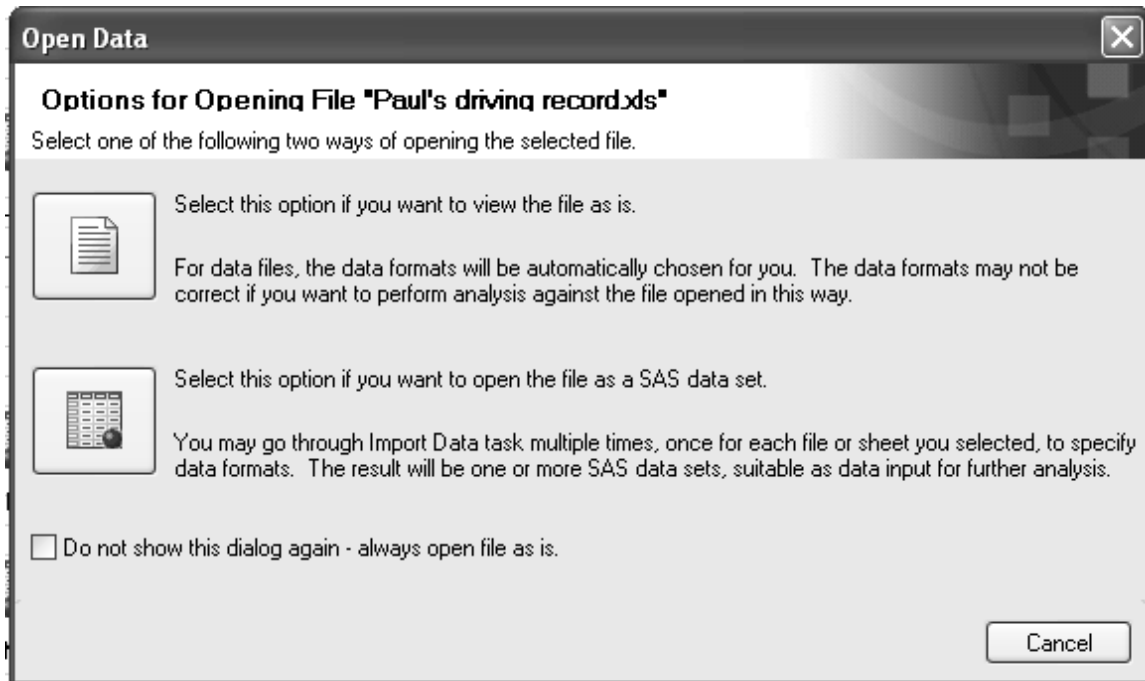


Figure 17: File open choices

Select the SAS data set option and you will open a dialog box with three areas you can set, 'Region to import', 'Column Options' and 'Results'. The 'Region to import' lets you select how many rows to import and from where. It also lets you specify a line as the column headings. 'Column Options' lets you specify the column names and formats

You can also specify where your SAS data set will go. The default is WORK.IMPW:

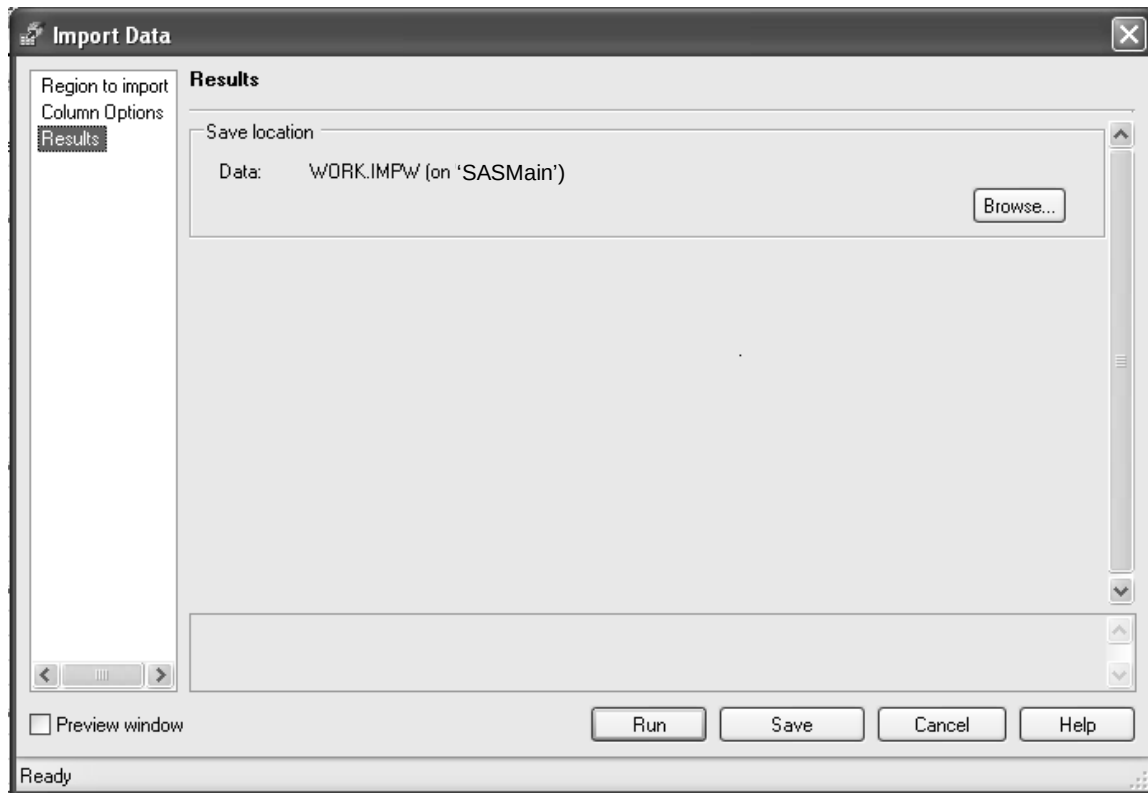


Figure 18: Results name

Once your spreadsheet is imported, the data set will open up whether or not you selected automatic opening of data when added to the project:

Project Designer		WORK.IMPW (read-only)				
	Date	Location of practice (eligibility)	Day or Night	Weather Conditions	Time spent	Initials of Adult
1	21OCT2007:12:00	highway	night	warm/fair	30 min.	P.H.
2	21OCT2007:12:00	residential	day	warm/sunny	10 min.	P.H.
3	26OCT2007:12:00	highway	night	rainy	15 min.	J.B.

Figure 19: After the import

VISUALIZING SAS DATASETS

You open a SAS dataset in much the same way as you opened existing code. The major difference is that you would usually want to open a library. Remember to have the automatic open data option turned off so you don't lock up your project.

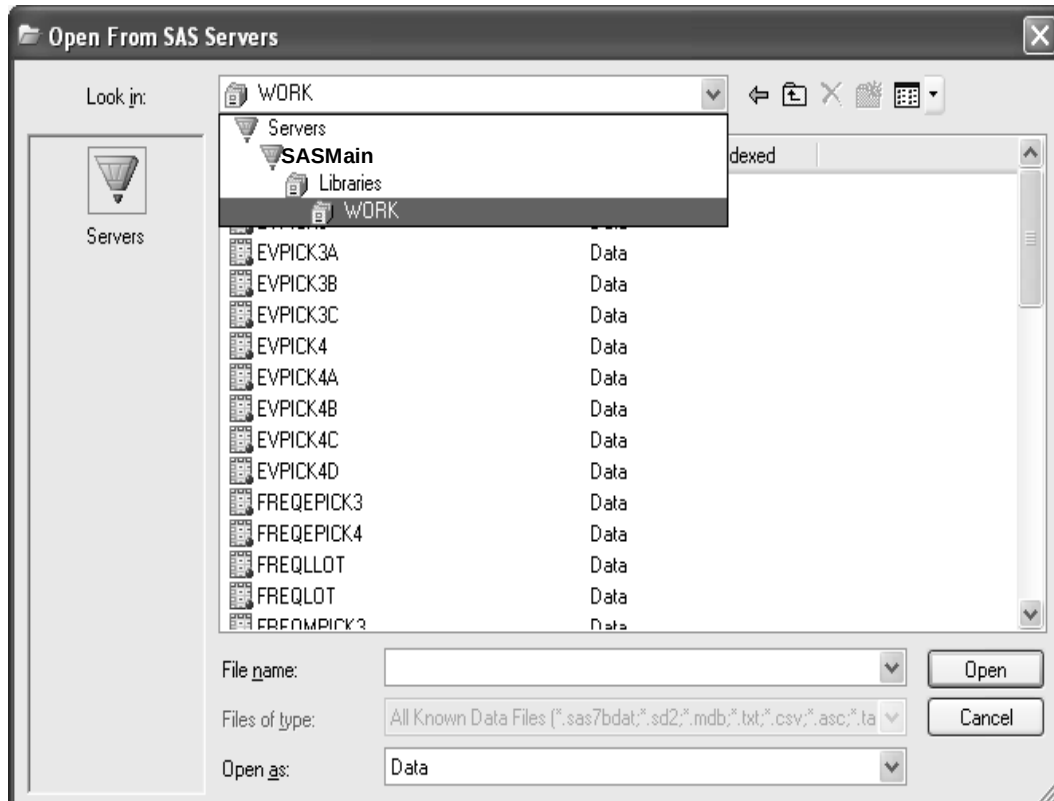


Figure 20: Adding a SAS dataset to your project

Selecting any of these will add the dataset to your project as an icon.

You may right-click on the data icon and look at the properties. There are two really useful categories, 'General' and 'Columns'. 'General' will give you the creation date, the number of columns and the number of rows. 'Columns' will give a list of all the columns, like a proc contents without the coding.

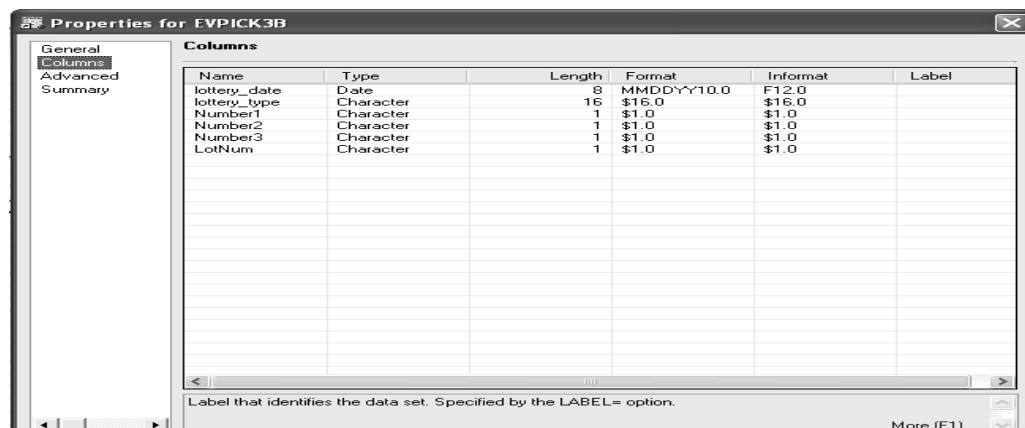


Figure 21: Variable list

Double-click on a dataset icon to open the table in read-only mode. Deselect the read-only option to edit. You will be asked if you want the data in update mode:

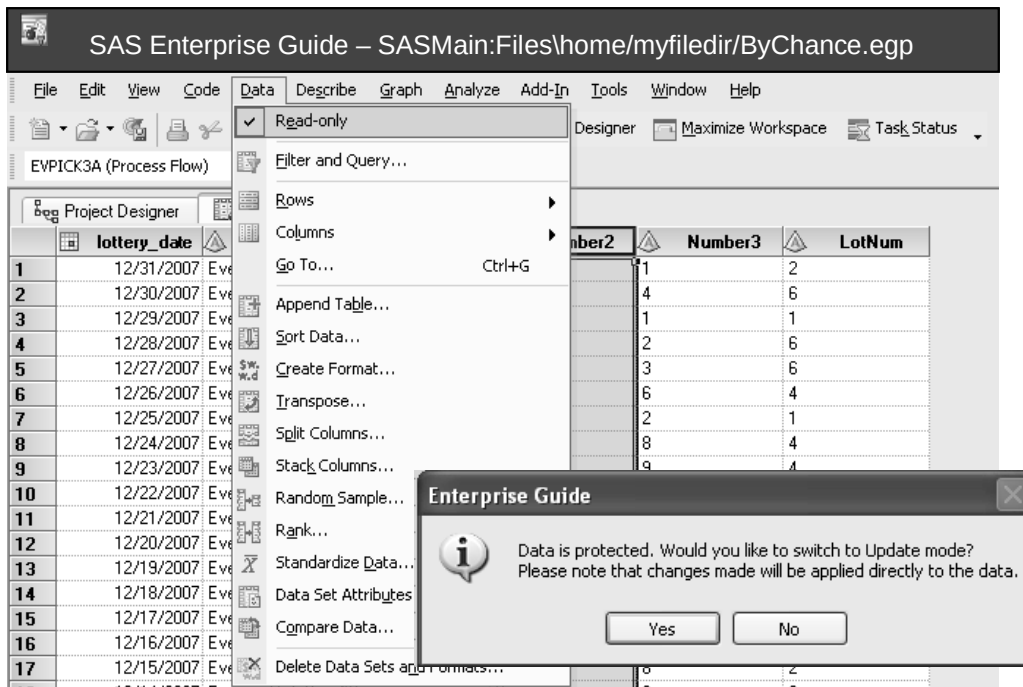


Figure 22: Editing a SAS dataset

You may copy and paste individual fields, columns or rows as well as type in new values. You will not be prompted for confirmation of changes and you may not be able to undo updates.

It is beyond the scope of this paper to go into any great detail about the many built in tasks. However, some understanding of the Query Builder would be profitable. It may be accessed from either the 'Data' toolbar or by right clicking on a dataset icon.

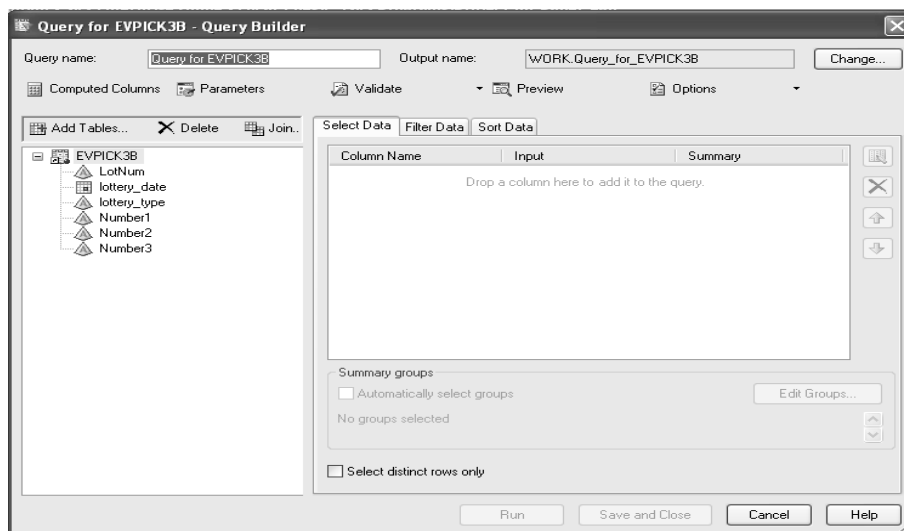


Figure 23: Query Builder screen

Once opened, you drag and drop the fields you want see in the results over to the Select Data tab.

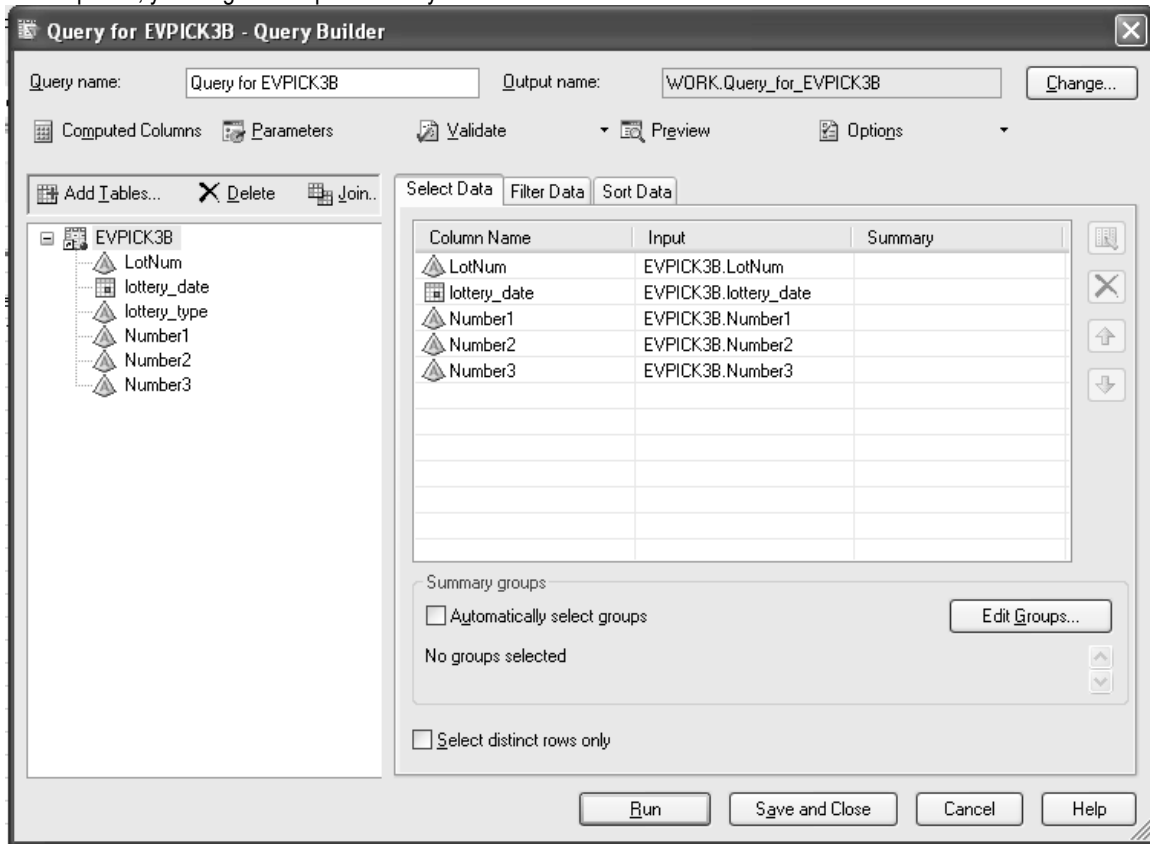


Figure 24: Select columns to be queried

You set your criteria in the 'Filter Data' tab, which may be edited:

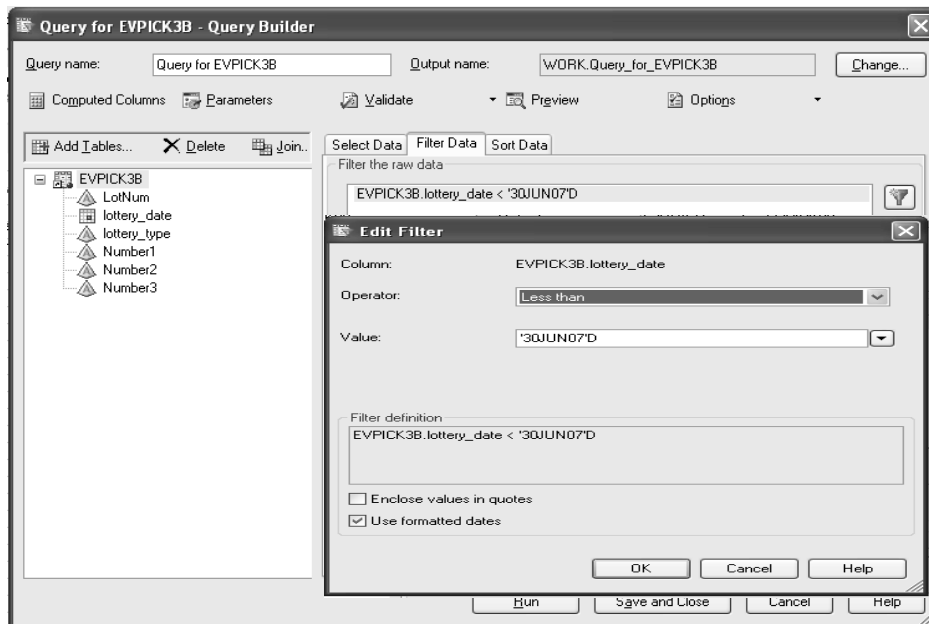


Figure 25: Select criteria

The results may be sorted as well:

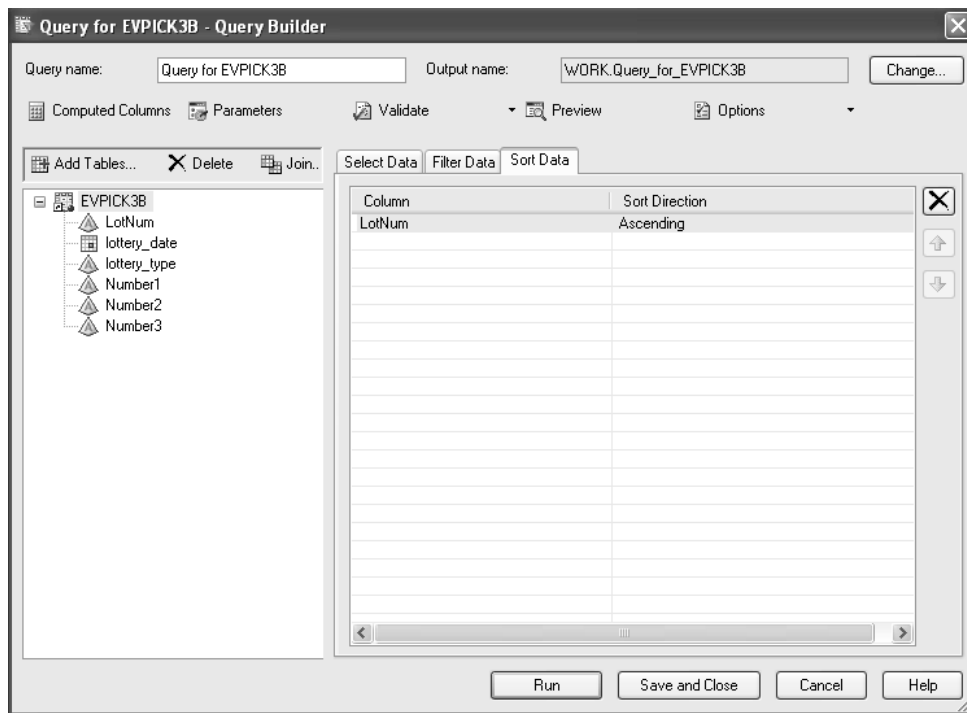


Figure 26: Sort selection

How this looks in Project Explorer:

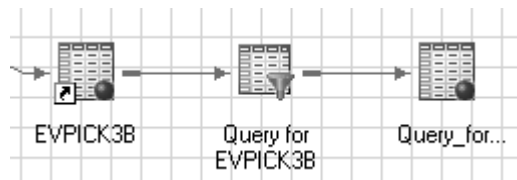


Figure 27: Project Explorer view of query results

The middle icon is the task and the icon furthest on the right designates the results in a SAS dataset. These can be queried again or used as input for other SAS procedures.

GETTING YOUR OUTPUT OUT, EVEN FROM CITRIX

You can generally send code, logs, reports and datasets as attachments to an email recipient from the 'File' toolbar or right-clicking on the icon and selecting 'send to email' recipient. When you do that, your email compose screen should come up automatically. However, if you are executing SAS from Citrix, you will probably need to export your output to your local drive first and send it from there. If your local drive isn't permanently mapped yet, your Citrix local drive, not your real local drive will come up first:



Figure 28: Your Citrix local drive

Find out how your local drives are truly mapped. In this example, they are mapped to 'V'. Type 'V:\' in file name and hit <ENTER>. Your 'C' drive should then appear and you can save your output there.

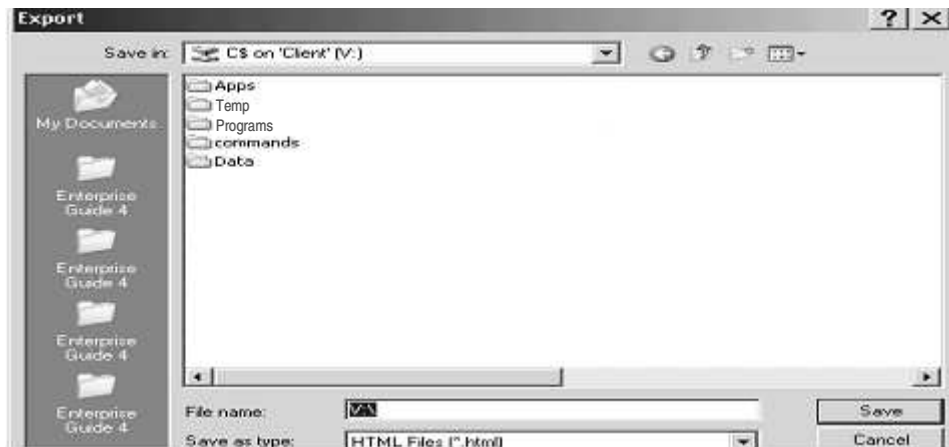


Figure 29: Your workstation's local drive as it appears in Citrix

You can save your output in various formats to your local drive but not in SAS format unless you have a local server.

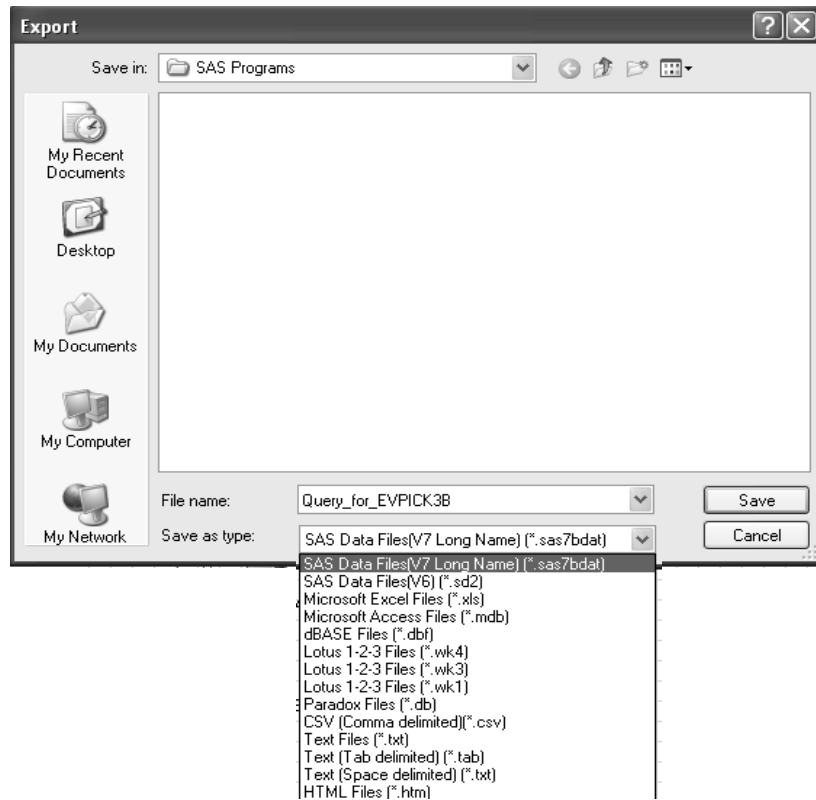


Figure 30: Saving your output

CONCLUSION

These tips should help in moving your code to SAS® Enterprise Guide®. Once there, you should try some of the built-in tasks on the toolbars for further analysis. As a starting point, 'Characterize Data' on the 'Data' toolbar is highly recommended. This feature does both proc means on numeric variables, including dates and proc freq on character values automatically to give you a good understanding of your data.

APPENDIX - COMMON INSTALLATION ISSUES

1. No system commands. This is the default. This means the X statement doesn't work either. Therefore you must go outside of SAS® EG to do things like change a dataset's security. It's hard to understand the logic here when most installations allow such commands in batch Unix. This setting must be changed the support people responsible for installation.
2. No email set up from SAS® EG. This means you can't set up an email step as part of the project but must do this either manually or code it if an email server is available.
3. Cannot use procs export or import to Microsoft Office even though proc setinit indicates that SAS/Access for Microsoft Office is installed. If you can't do this executing your program straight from the Unix prompt, it's probably not installed on Unix. See your support people.

ACKNOWLEDGEMENTS

Thank you, Joe and Paul Butkovich for reviewing this paper and presentation

CONTACT INFORMATION

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

Patricia Hettinger

Oakbrook, IL 60523

Phone: 630-847-2728

Email: hettipatra@yahoo.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.