SAS® IT Resource Management Forecasting

The Power to Know.

A SAS White Paper

Table of Contents

Introduction to SAS IT Resource Management Forecasting **1**

Getting Started with the SAS Enterprise Guide Administrator **1**

Using SAS Enterprise Guide **6**

Basic Forecasting Example 1 ..... **6**

Forecasting Example 2 ..... **10**

How to Create a Plot with SAS Enterprise Guide **13**

Additional Considerations: Forecasting with a PDB **20**

Conclusion **20**

Why SAS? **20**

Introduction to SAS IT Resource Management Forecasting

SAS Enterprise Guide is an application written by SAS that runs under the Windows operating environment. SAS Enterprise Guide provides a point-and-click interface that generates SAS code and produces results from licensed SAS products. Many types of data residing on a variety of computer servers can be accessed and analyzed by using SAS Enterprise Guide, which is provided with SAS IT Resource Management. This paper describes how to set up SAS Enterprise Guide and use it to analyze and forecast data that resides in a SAS Performance Database (PDB). Information about the Time Series Forecasting System (TSFS) as another way to interactively analyze time series data is also described.

Because familiarity with SAS IT Resource Management administration is presumed, the management of the PDB is not discussed in this paper.

Before running a forecast, you have to know how much data you have, for example, two years of monthly data is required to determine seasonality in a time series. You also have to know what level of summarization you need, how the data will be aggregated, and what metrics and statistics you want to use. You might want to forecast activity according to a specific work shift, which will require additional data preparation. You might want to forecast server activity by using one number per month per machine. Or, you might want to create a 24-hour time series for each machine where each hour is forecast month to month.

Data in the PDB might need to be re-classified or re-summarized prior to forecasting. PROC EXPAND will have to be used to prepare data by interpolating for missing values and correcting for different numbers of days in a period.

All of these functions are easily performed by using SAS Enterprise Guide, which is capable of quickly creating forecasts for hundreds of time series in a single step. SAS provides many exciting and easy-to-use analytical capabilities for both novice and experienced users.

Getting Started with the SAS Enterprise Guide Administrator

Prior to using the SAS Enterprise Guide application with a PDB, the Enterprise Guide Administrator component must be used to make the data available for any analysis tasks that will be performed. The Enterprise Guide Administrator interface enables you to specify a list of servers and data libraries that you want to access. This interface is launched from the start menu on a PC.

SAS IT Resource Management PDB's can exist under Windows, z/OS, UNIX, and LINUX operating environments. You can access data that resides on any of these servers by using SAS Enterprise Guide and adding the appropriate server and library definitions. To access a PDB that exists on the local machine's Windows environment, the "local server" definition is used.

With SAS Enterprise Guide installed on your PC, the Enterprise Guide Administrator is activated by clicking the Windows shortcut on your desktop to invoke this application. Using this Administrator, you can specify connections to remote servers such as z/OS and UNIX servers where your data resides. Libraries and binders (groups of libraries) can also be defined.

The "local server" definition refers to the local PC where SAS Enterprise Guide is installed. You can specify a list of SAS libraries and find other files on disk drives that you can make available for SAS Enterprise Guide to use.

There are two ways to access PDB data. The first way is to define a SAS library that refers to a specific table level in a PDB such as Day or Month. After this library is defined, the SAS Enterprise Guide client will be able to access the SAS data in the library for analysis and forecasting.

The second way to access PDB data is to have the SAS Enterprise Guide Administrator server definition execute a %CPSTART macro on your behalf and allocate read-only access to the PDB that you select. To do this, you first create a file on the server where the PDB resides. This file contains the %CPSTART macro call for the selected PDB in batch mode. Then, in the SAS Enterprise Guide Administrator specify an INCLUDE statement in the server definition that points to this code. Later, when SAS Enterprise Guide opens the PDB, all the libraries are allocated for you, and you will be able to build your forecasts. More details about this are forthcoming.

With the local server selected, from the “File” menu, select “New” to open a window that contains the icons shown in Figure 1.

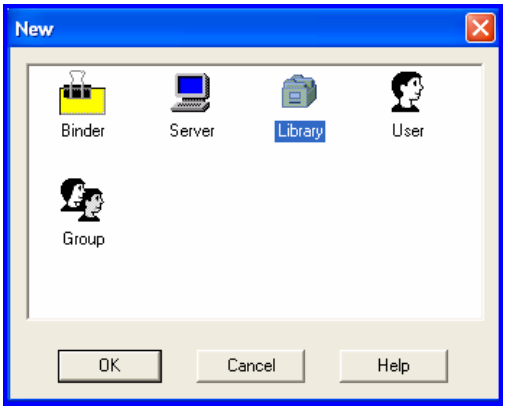


Figure 1. Adding a New Component in the SAS Enterprise Guide Administrator

PDB Access Technique 1: By selecting the “Library” icon, you can enter the name and description of the specific PDB level that you want to work with. The interface is intuitive and takes you through the necessary steps for adding libraries. Normally, SAS IT Resource Management uses a %CPSTART macro to initialize the environment to access PDB data that contains both views and data. Use of SAS Enterprise Guide libraries makes %CPSTART unnecessary, but you must name the library appropriately, otherwise the view will not work. For example, if you want to use the Day level of a PDB, be sure that you name the library DAY in the Enterprise Guide Administrator. This will ensure that all table components (multiple data sets and formula variables within the view) are available for your analysis work.

Figure 2 shows the Enterprise Guide Administrator with the SASUSER and WORK libraries as well as the additional libraries DAY and MONTH defined for the local server.

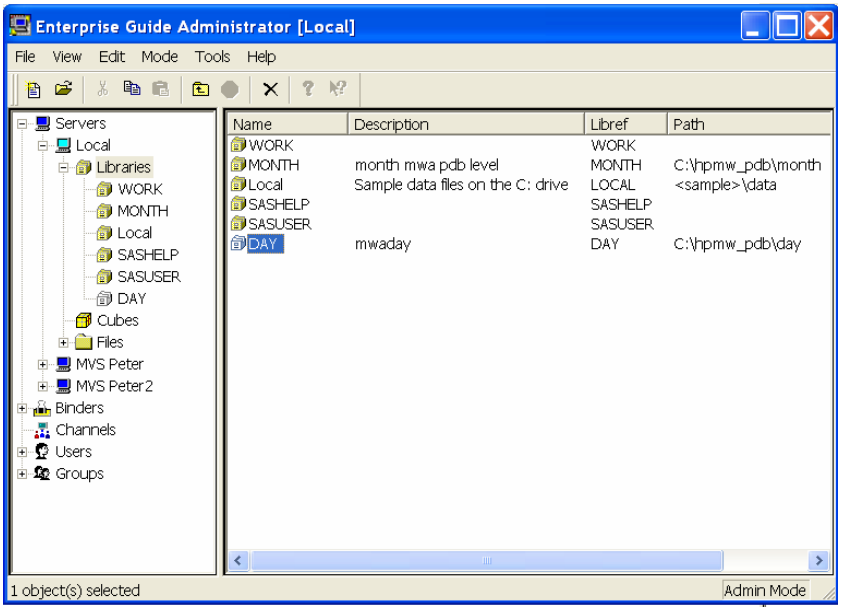


Figure 2. Local Server Library Definitions

PDB Access Technique 2: When creating a server definition in the SAS Enterprise Guide Administrator, you can automatically issue a %CPSTART for an entire PDB by using the OPTIONS tab and specifying the location of the %CPSTART code on the server via a %INCLUDE statement as shown in Figure 3.

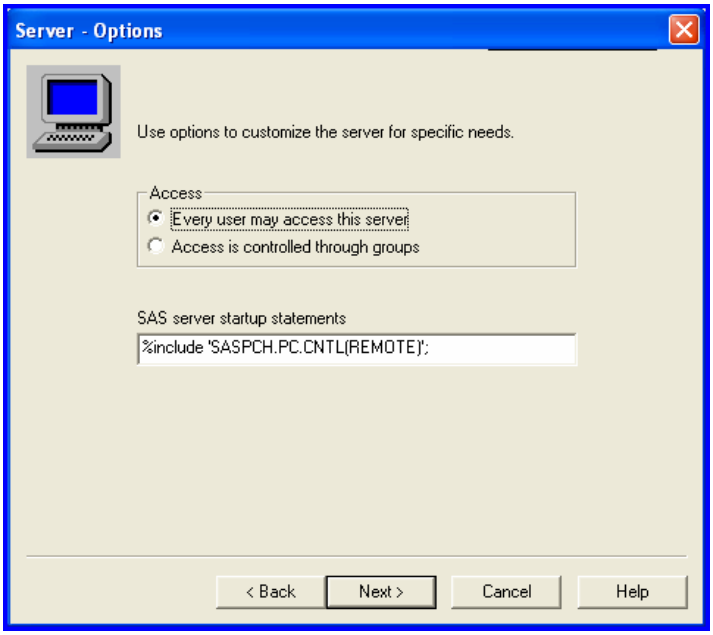


Figure 3. Server Options with Location of Code Containing %CPSTART Command

The code referred to in Figure 3 is shown in Figure 4. This is an example of how to activate a PDB under z/OS.

```
EDIT       SASPCH.PC.CNTL(REMOTE) - 01.02
Command ===> _____________________________________________
****** ***************************** Top of Data ***
000001 %CPSTART(MODE=BATCH,
000002          PDB=CPD.ITSV25.MVSGPP.PDBDEMO.,
000003          MXGSRC=('CPD.MXGCPE.SOURCLIB'
000004                  'SSD.MXGPROD.SOURCLIB'),
000005          MXGLIB=SSD.MXGPROD.LIBRARY,
000006          DISP=SHR,
000007          SHARE=N/A,
000008          ROOT=CPD.TS0802M1.
000009          );
****** ***************************** Bottom of Data
```

Figure 4. % CPSTART to Allocate a PDB That Resides on a z/ OS Server

When the code in Figure 4 and the %INCLUDE statement in Figure 3 are defined in a SAS Enterprise Guide server definition, all libraries in the PDB are available to the SAS Enterprise Guide client. Notice that in the right-most frame of the SAS Enterprise Guide client in Figure 5, all levels of the PDB under z/OS are available. In the center frame, you can see a portion of the XRMFINT table as well as a line plot and a bar chart that were created with the SAS Enterprise Guide project that is defined in the top-left frame. This is discussed in more detail later in the paper.

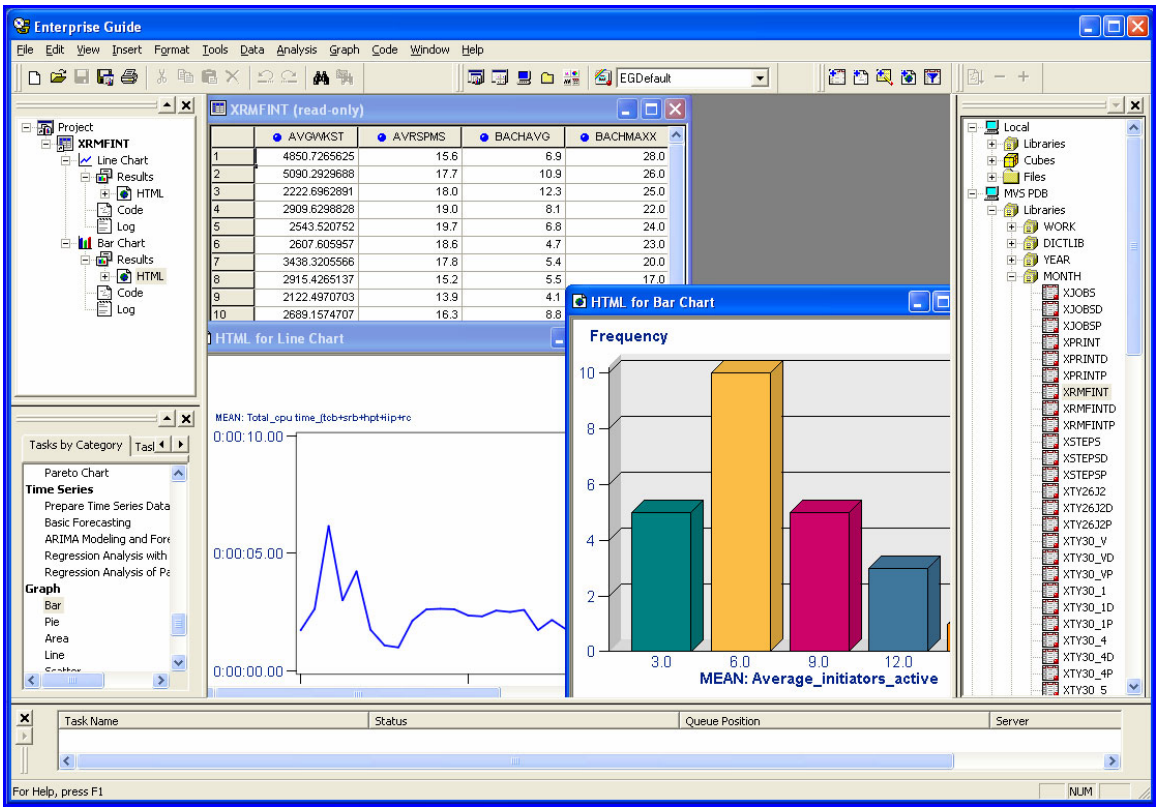


Figure 5. PDB Libraries Made Available via an Included %CPSTART Command

In the New window (Figure 1), notice there is also an icon called “Server”. This is how the administrator defines connectivity to a server other than the local PC. Server definitions may be specified to access data on z/OS operating environments or UNIX operating environments.

To access a z/OS or a UNIX operating environment, use the IOM (Integrated Object Model) in SAS IT. The server must have IOM installed and configured by your systems administrator, and then the SAS Enterprise Guide user will be able to define the properties of a connection, including a port address and host name. The steps for setting up IOM on the server are not included in this paper.

If you want to use PDB access technique 1, the library that you want to work with on z/OS will still have to be defined as a library by using the SAS Enterprise Guide Administrator in the same way as it was used in the Windows example. Instead of using the “local server” definition, the definition for the z/OS server should be selected.

Bound Libraries (Figure 6) are optional definitions that enable the user of SAS Enterprise Guide to enter a naming convention pattern that will dynamically list the data sets that match the pattern. If you specify the high-level qualifier and pattern data set name, you will be permitted to browse through all data sets on the server that match the pattern.

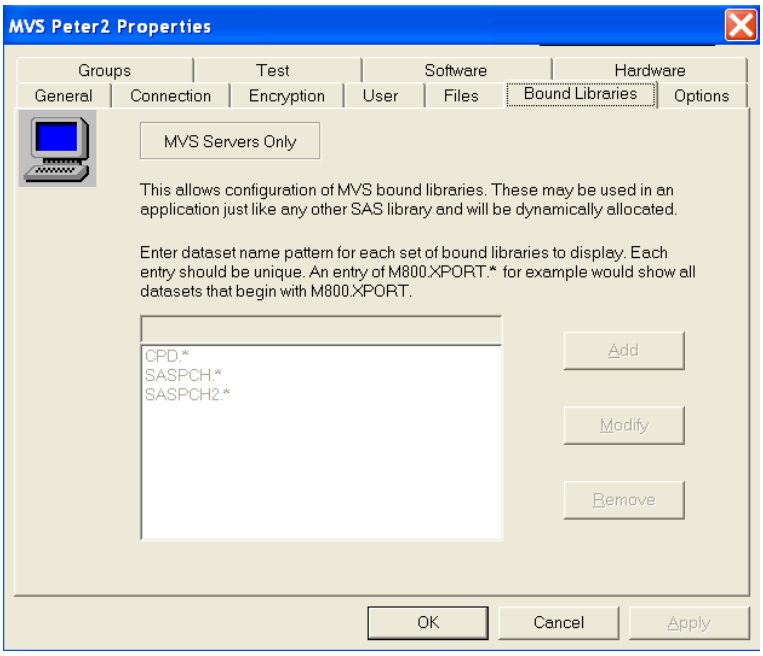


Figure 6. Bound Libraries and Property Tabs

Figure 7 shows the DAY tables for a PDB, which resides on a z/OS operating environment. Notice that both the SAS view and the data component are visible. For example, XJOBS is the SAS view, which also contains formula variables, while XJOBSD is only the data component that the view uses. Remember to always use the view when reporting against PDB tables so that all the formula variables will be available.

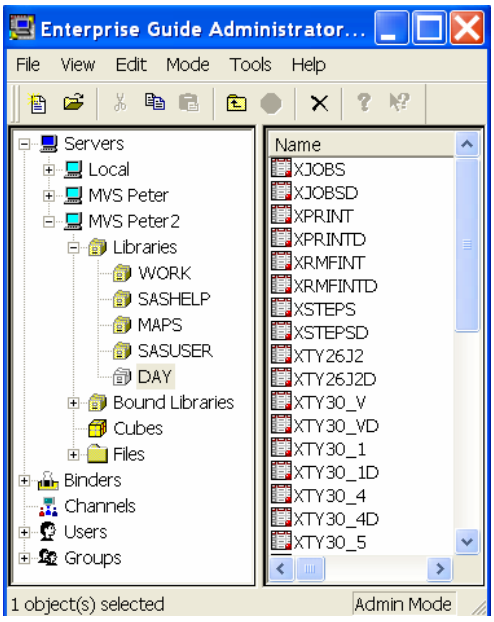


Figure 7. z/ OS PDB Tables in Enterprise Guide

Using SAS Enterprise Guide

After the server library definition is complete, you can start SAS Enterprise Guide on your PC to create and name a new Project. You add data to your analysis project by clicking the appropriate server icon. In this first example, the PDB is on the Windows operating environment so the “local server” definition is used. By clicking the plus sign (+) next to Local, you can expand the “Libraries” list, which shows the SAS libraries that were defined through the administrator interface. By further expanding the Day level, you can see all the tables within your PDB. Data is available for analysis by double clicking the data in the server window.

Always use the SAS view when exploring data with SAS Enterprise Guide. Notice in the top-left frame of Figure 8 that the UNIX application data PCSAPPD and PCSAPP from the Measureware table has been made available to the project. The first is the SAS IT Resource Management Table “Data Component”; the second is the view. The view contains formula variables such as CPUAPPL and DATE that do not appear in the Data Component. It is best to always use the view to make sure that all variables are available. You will be able to do a “File -> Save” from the pull-down menu to keep any analysis that is performed by using SAS Enterprise Guide.

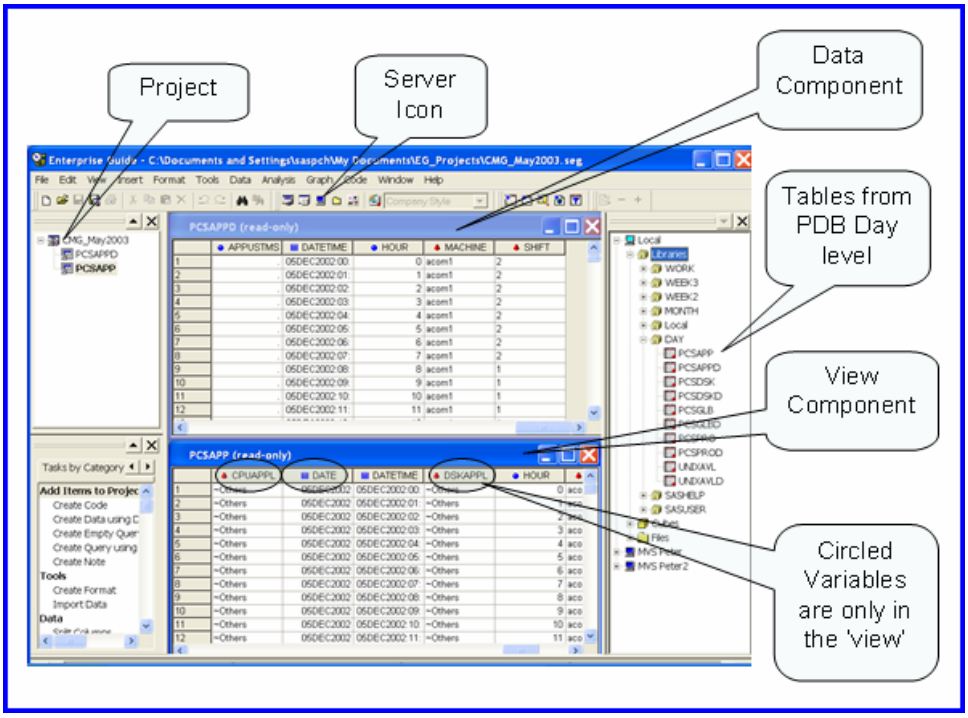


Figure 8. SAS Enterprise Guide Project

Basic Forecasting Example 1

Premise: Use the Month level of the PCSGLB table that contains 24 months of data over 4 shifts. You want to analyze the prime shift CPU, memory, network packet rate, and disk I/O rate over time and forecast these variables into the future. By selecting only the prime shift, activity that occurs on weekends or holidays will be excluded from the data. Remember that you must decide what you want to forecast. This example creates a separate forecast for each hour, over all the months in the PDB. Each hour represents a time series. There are many other ways to create forecasts.

Note: Forecasting must use a percent or a rate variable, not a count unless count is weighted or normalized correctly.

Here are the three steps to produce a basic forecast with SAS Enterprise Guide:

1. Subset the data to exclude any data that is not relevant to your forecast such as other machines or holiday shifts.

This is done by using the query icon in SAS Enterprise Guide and selecting the specific machine (s) and shift as shown in Figure 9. The subset data is not yet in the order that is necessary to build a time series and an appropriate forecast.

If you want a specific machine and shift at the Month Summary level, select only the data for that machine. In this example, MACHINE ='shiva' and SHIFT ='1' are selected from the Month level of the PDB table. This will also reduce the amount of data that is read from the table.

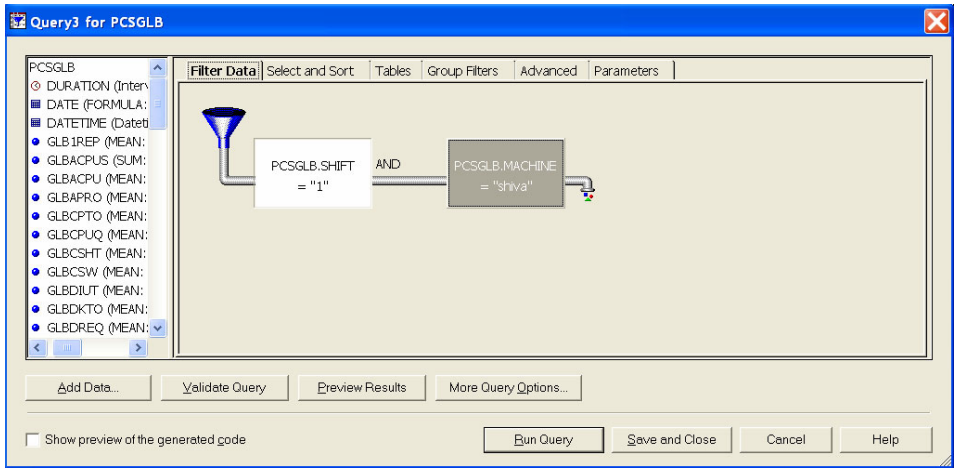


Figure 9. Data Filter

Figure 10 shows how the data looks after the subset operation. Notice that the data is sorted in the PDB according to month and hour because of the classification variables MACHINE, DATETIME, and HOUR.

Results of Query3 for PCSGLB (read-only)

	DURATION	DATE	DATETIME	GLB1REP	GLBACPUS
1	15:58:29.00	01JUL2001	01JUL2001:08:00:0	0	
2	15:58:37.00	01JUL2001	01JUL2001:09:00:0	0.0003638924	
3	15:58:34.00	01JUL2001	01JUL2001:10:00:0	0.0004165942	
4	16:05:12.00	01JUL2001	01JUL2001:11:00:0	0.0001554082	
5	15:52:04.00	01JUL2001	01JUL2001:12:00:0	0	
6	15:58:36.00	01JUL2001	01JUL2001:13:00:0	0.0046281035	
7	15:58:41.00	01JUL2001	01JUL2001:14:00:0	0.0122558718	
8	15:58:30.00	01JUL2001	01JUL2001:15:00:0	0.0013540254	
9	15:58:36.00	01JUL2001	01JUL2001:16:00:0	0.0005718409	
10	22:57:55.00	01AUG2001	01AUG2001:08:00:	0.000797339	

Figure 10. Resulting Subset Data

2. Prepare the time series data.

In the Month level of the PDB, each hour represents a time series that contains the requested analysis variables. There is a time series for summarized data for 8:00 a.m.– 9:00 a.m. as it varies month-to-month and 9:00 a.m.– 10:00 a.m. month-to-month and so on.

From the data subset, select DATE as the time ID variable, select the analysis variables that you want as the time series variables, and group the data by HOUR. The SAS Enterprise Guide user interface with the appropriate selections made is shown in Figure 11.

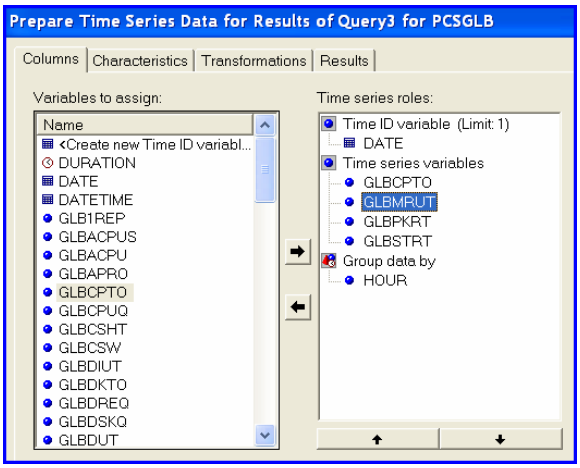


Figure 11. Prepare Time Series Data

After the data has been placed into time series format, it looks like Figure 12. Compare this to the results of the subset data in Figure 10 to see the difference in data that is necessary to produce a forecast. Notice that the data contains only the time-series variables (GLBCPTO, GLBMRUT, etc.) that you want to forecast. These variables are sorted by HOUR and MONTH.

TSD14026 (read-only)

	HOUR	DATE	GLBCPTO	GLBMRUT	GLBPKRT	GLBSTRT
1	8	01JUL2001	33.70%	51.72%	311.71312316	1.1741449164
2	8	01AUG2001	50.38%	51.83%	213.74019958	2.1036722105
3	8	01SEP2001	54.98%	53.94%	146.01449773	1.1975995287
4	8	01OCT2001	48.25%	54.16%	126.90004475	1.9584009578
5	8	01NOV2001	52.76%	51.14%	221.32891479	1.273650096
6	8	01DEC2001	49.02%	50.35%	158.31466285	1.2141287415
7	8	01JAN2002	33.97%	49.67%	144.08281179	0.9481790547
8	8	01FEB2002	54.92%	51.25%	311.44965458	1.1532324622
9	8	01MAR2002	49.66%	53.10%	100.81275211	1.4175202381
10	8	01APR2002	46.60%	53.77%	13.971484212	1.30396439
11	8	01MAY2002	0.53%	49.25%	6.9291249404	0.4994098712
12	8	01JUN2002	17.27%	49.50%	258.39376912	1.9107482615
13	8	01JUL2002	59.95%	53.13%	265.21491551	5.3125545786
14	8	01AUG2002	67.81%	55.07%	321.65802861	4.3525039746
15	8	01SEP2002	53.14%	56.61%	183.37081048	3.9489454998
16	8	01OCT2002	58.82%	52.11%	271.57522487	4.0419012545
17	8	01NOV2002	45.28%	50.31%	160.69835273	3.8359587085
18	8	01DEC2002	41.56%	49.87%	166.47174899	3.5423979111
19	9	01JUL2001	32.50%	51.56%	206.74607507	3.6131526331

Figure 12. Resulting Data

3) Use SAS Enterprise Guide to build and run the basic forecast by dragging the variables that you want from the window at the left and dropping them under the appropriate roles in the window at the right, as shown in Figure 13. After filling in all the options on the tabs, you can run the basic forecast by ensuring that the “Run task now” box is checked and then clicking the ”Finish” button. You can also view the SAS code that was generated by the dialog.

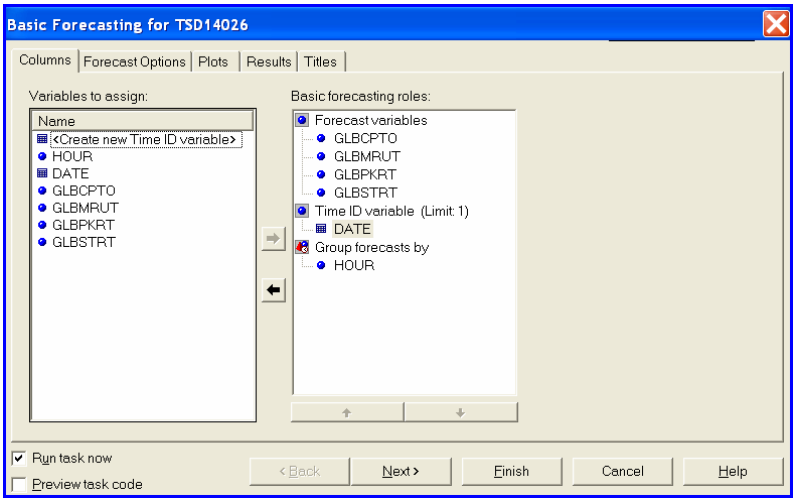


Figure 13. Building a Forecast

Figure 14 shows the results of one of the time series forecasts that was produced for 9:00 a.m. of the network packet rate for the past 24 months, forecast one year into the future. If SAS is able to identify a pattern or seasonality in the existing data, then the forecast takes this pattern into account. Notice that the forecast line at the right is not a straight or a curved line, but it follows a 'zigzag' pattern, which was identified in the historic data. If no pattern was found, then the forecast line would be a straight or a curved line. Upper and lower 95% confidence limits for the forecast are shown on the plot.

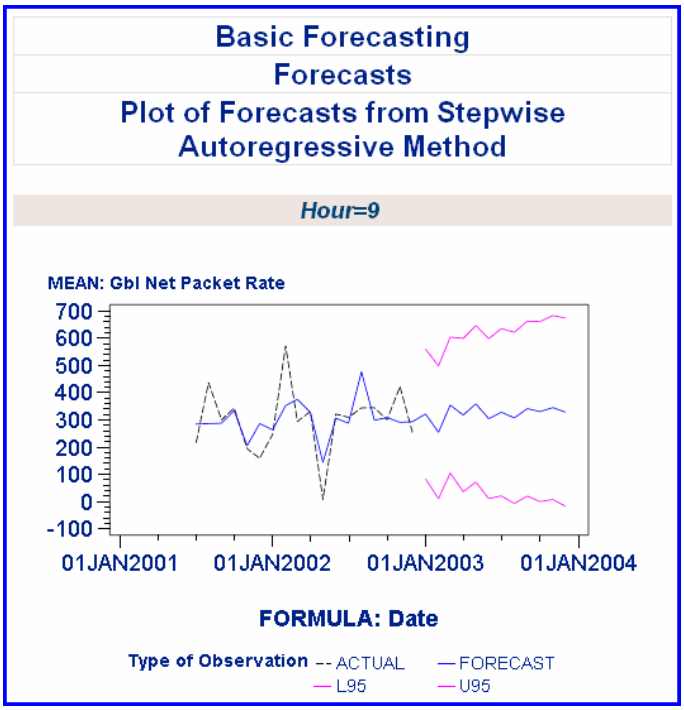


Figure 14. Forecast Plot

Forecasting Example 2

Suppose you want to produce a forecast that uses a single number per month to represent prime-shift CPU usage as the time series, instead of using a separate number for each hour for each month. In this case, the data from the Month level in the PDB must be re-summarized to provide a single number for each month if the PDB has HOUR as a class variable. You can further subset the data according to shift.

Using the data in Figure 10 in Forecasting Example 1, you can use the SAS Enterprise Guide task ”Descriptive summary statistics” to create the summarized data. Figure 15 shows information in the “Columns” tab completely filled out to re-classify data according to MACHINE and DATE.

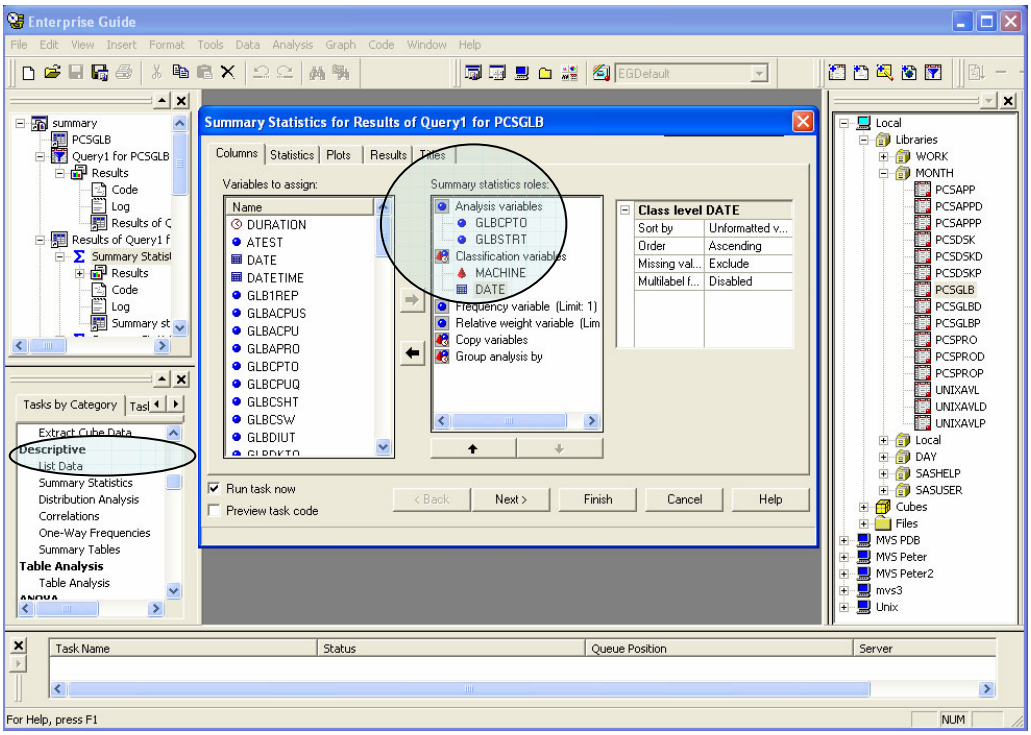


Figure15. Create Summary Statistics

Select the ”Results” tab to save the output data table in a SAS library, as shown in Figure 16. If you don’t specify that SAS Enterprise Guide should save the results, by default, you will just see the tabular html or PDF output. The other tabs enable you to specify the statistics that you want, the plots to be created and titles.

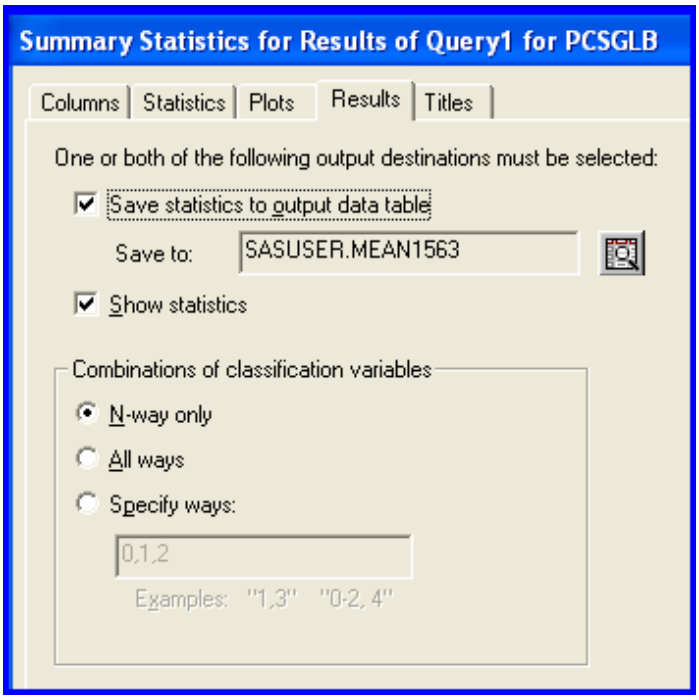


Figure 16. The Results Tab

Next, perform the “Prepare time series data” task similarly to the way it was done in Example 1 except without specifying an hourly summary against the data table that was just created. The resulting data is in Figure 17. Notice that now, instead of having a number for each hour for each month, you have only one number that represents the mean of all hours for each month. Figure 18 shows the resulting forecast.

Remember that you must decide what you want to forecast and prepare the data appropriately. There are cases when selecting the maximum value might result in a graph of a straight line at 100% for every month if there are usually peaks of 100% in the workload. This would not be useful in creating a meaningful forecast. Selecting the mean statistic would probably reveal a more useful pattern. By using SAS Enterprise Guide, it is easy to analyze your data and quickly create forecasts that make sense in your environment.

SUBSET (read-only)

	DATE	MGLBCPTO	MGLBSTRT
2	01AUG2001	6.34%	2.10952869
3	01SEP2001	7.13%	2.3294543521
4	01OCT2001	4.77%	1.7798828773
5	01NOV2001	4.23%	0.85835092
6	01DEC2001	6.78%	1.964015609
7	01JAN2002	5.36%	1.2812633946
8	01FEB2002	8.05%	1.3584587903
9	01MAR2002	7.78%	2.7711774157
10	01APR2002	13.07%	2.086893309
11	01MAY2002	9.15%	1.7936046777
12	01JUN2002	22.98%	18.000855916
13	01JUL2002	7.80%	1.4694828044
14	01AUG2002	11.64%	1.53395181
15	01SEP2002	9.37%	1.8268727642
16	01OCT2002	6.04%	2.9146439892
17	01NOV2002	7.87%	4.5793106171
18	01DEC2002	8.82%	3.9563157241
19	01JAN2003	12.34%	2.1337981224
20	01FEB2003	13.31%	1.9957036209
21	01MAR2003	24.78%	3.2783097628
22	01APR2003	13.01%	1.8963045517
23	01MAY2003	15.44%	3.564268811
24	01JUN2003	15.62%	2.5971323579
25	01JUL2003	12.33%	2.7468656335

Figure 17. Resulting Subset Data

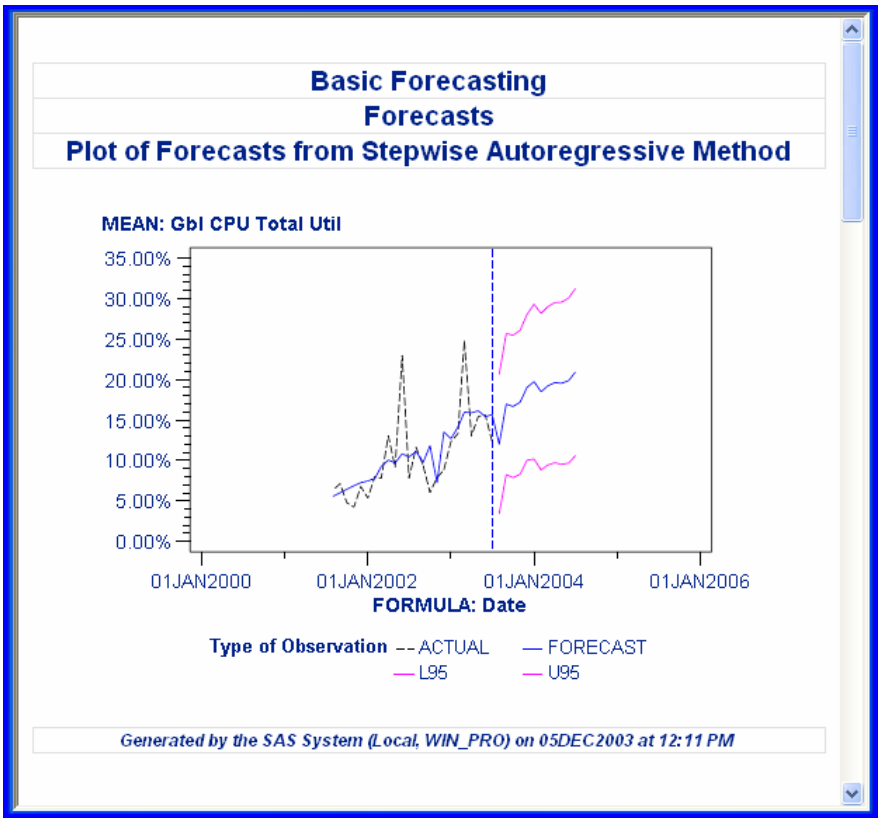


Figure 18. The Forecast

In SAS Enterprise Guide, select a “Line Plot” task from the available chart types shown on the “Chart Gallery” tab. (See Figure 19 for examples of some chart types.)

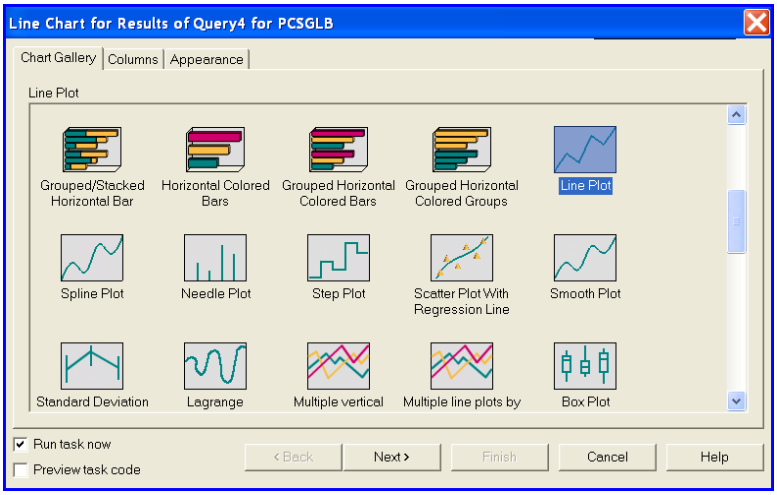


Figure 19. Line Plot Selections

Then, click the “Columns” tab, and drag and drop the variables for the 2-y plot from the “Columns to assign” list box on the left to the “Line Roles” list box on the right. (See Figure 20.)

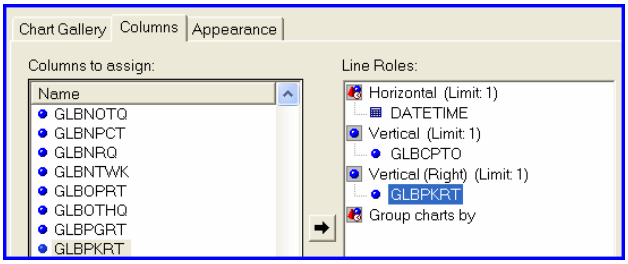


Figure 20. Create a 2-y Plot

If you run this task right now, SAS Enterprise Guide would try to plot data for all servers in the warehouse. Because you are only interested in LOGIN021, you can modify the code that was generated by SAS Enterprise Guide to select only this server. (Alternatively, you can choose to change the filter in SAS Enterprise Guide to do the subsetting.)

```
FOOTNOTE "Generated by the SAS System
data work.plot;
set sasuser.qury4623;
if machine="login021";
PROC gplot DATA=work.plot
     ;
     plot GLBCPTO*DATETIME /
     LEGEND=LEGEND1
     VAXIS=AXIS1
     HAXIS=AXIS2
     FRAME
```

Figure 21. Code to Subset Server

Figure 22 shows the resulting graph of the LOGIN021 server, which does not have a correlation between CPU and Packet rate.

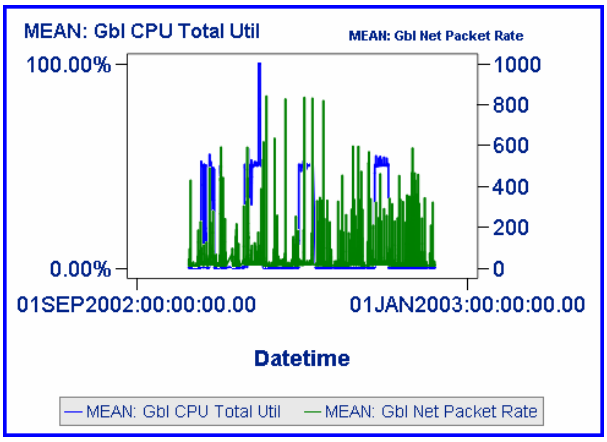


Figure 22. 2- y Plot

The 24-hour summary of this data (Figure 23) looks like there might be a correlation between 9 and 5, so you might want to re-run the analysis to better model the correlation effect of these variables during the prime shift.

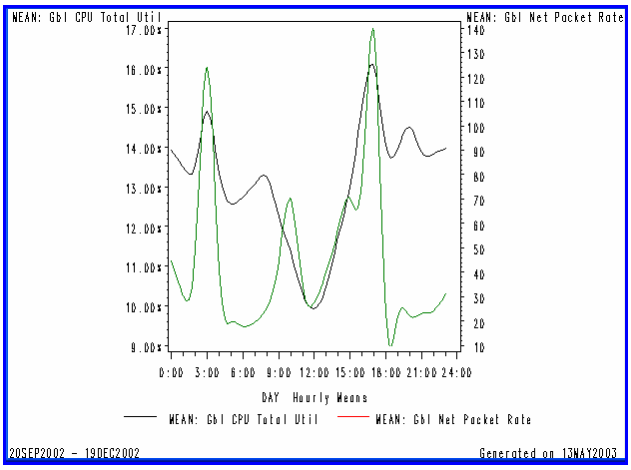


Figure 23. Summary Plot

Using the Time Series Forecasting System (TSFS)

SAS IT Resource Management also provides the TSFS capability as another way to model time series data and produce forecast models and forecast graphs. TSFS is provided in addition to SAS Enterprise Guide. TSFS is activated from the SAS IT Resource Manager user interface. In the SAS Editor window, from the ”Solutions” menu select “Analysis” and then “Time Series Forecasting System.”

The PDB data will have to be prepared for use by TSFS. The data preparation can be performed by using SAS Enterprise Guide, or the data can be prepared directly by executing a DATA step in a Program Editor window. In this example, the Program Editor window is used to prepare the data.

The SAS IT Resource Management PC GUI, connected to a z/OS PBD Month-level XTY70 table, is the source of the data for this analysis exercise. The first thing you should do is subset the data for your system, and choose only the variables that you want from the 372 variables that actually reside in this PDB.

The following code, which is submitted from the PC, reads the monthly summary level from the z/OS PDB table ”month.xty70” and creates the subset of data that is used in the TSFS exercise. This code also excludes holiday shifts

and adds the DATE and TIME variables to the copy of the data that will be held in the SASUSER data set on the PC. Because the data is already highly summarized at the monthly reduction level and only a few variables are being kept, the resulting data set is small, even though it represents 24 months of data.

```
/* select the variables that you want from the PDB table */

data sasuser.mvstm (keep=datetime iorate machine shift pccpuby tsomax_x stcmax_x date
hour);
set month.xty70;
if shift ne '0';
format date date9.;
format hour time5.;
date = datepart(datetime);
hour = timepart(datetime);
run;
```

Figure 24 shows some of the resulting analysis variables.

	Datetime	MEAN: Io activity rate	System id	MEAN: All cpus percent busy (dispatched)	Operations shift	MAX: Maximum stc users	MAX: Maximum tso users
1	01JUL2001:00:00:00.00	726.716	PRC4	50.37	2	202.0	28.0
2	01JUL2001:01:00:00.00	899.514	PRC4	54.29	2	201.0	27.0
3	01JUL2001:02:00:00.00	933.441	PRC4	53.51	2	201.0	25.0
4	01JUL2001:03:00:00.00	417.938	PRC4	46.38	2	203.0	29.0
5	01JUL2001:04:00:00.00	690.396	PRC4	70.26	2	200.0	29.0
6	01JUL2001:05:00:00.00	683.865	PRC4	68.51	2	199.0	31.0
7	01JUL2001:06:00:00.00	650.642	PRC4	61.82	2	200.0	36.0
8	01JUL2001:07:00:00.00	499.891	PRC4	41.70	2	202.0	67.0
9	01JUL2001:08:00:00.00	629.188	PRC4	44.97	2	201.0	178.0
10	01JUL2001:09:00:00.00	576.565	PRC4	50.80	1	190.0	303.0
11	01JUL2001:10:00:00.00	708.130	PRC4	63.56	1	190.0	361.0
12	01JUL2001:11:00:00.00	672.625	PRC4	56.34	1	191.0	380.0
13	01JUL2001:12:00:00.00	622.999	PRC4	46.63	1	190.0	394.0
14	01JUL2001:13:00:00.00	637.680	PRC4	51.76	1	189.0	403.0
15	01JUL2001:14:00:00.00	672.818	PRC4	53.93	1	190.0	407.0
16	01JUL2001:15:00:00.00	648.918	PRC4	55.00	1	194.0	400.0
17	01JUL2001:16:00:00.00	636.857	PRC4	50.63	1	196.0	366.0
18	01JUL2001:17:00:00.00	502.182	PRC4	43.98	2	192.0	268.0

Figure 24. z/OS Subset Data

The next step is to transpose this data into the time series data format in which each hour represents a time series. This will give a time series for summarized data for 9:00 a.m. –10:00 a.m. as it varies month-to-month, and 10:00 a.m.–11:00 a.m. month-to-month, and so on. In this example, the HOUR variables are renamed with the prefix “Hour”, so that the first hour, ”Hour1“, represents 00:00–01:00 a.m. in the prepared data set. You can use the following code to accomplish this.

```
/* create the time series for the monthly data where each hour will be a series */

proc transpose data=sasuser.mvstm out=sasuser.mvstrans
prefix = hour;
by date ;
var pccpuby;
run;
```

The resulting transposed data now looks like Figure 25.

VIEWTABLE: Sasuser.Mvstrans

	DATE	NAME OF FORMER VARIABLE	LABEL OF FORMER VARIABLE	HOUR1	HOUR2	HOUR3	HOUR4	HOUR5	HOUR6	HO
1	01JUL2001	PCCPUBY	MEAN: All cpus percent busy (dispatched)	50.37	54.29	53.51	46.38	70.26	68.51	6
2	01AUG2001	PCCPUBY	MEAN: All cpus percent busy (dispatched)	58.49	62.51	57.91	49.85	73.68	70.52	6
3	01SEP2001	PCCPUBY	MEAN: All cpus percent busy (dispatched)	59.41	57.81	53.60	44.06	73.25	67.74	6
4	01OCT2001	PCCPUBY	MEAN: All cpus percent busy (dispatched)	57.40	54.10	50.04	41.42	70.96	71.15	6
5	01NOV2001	PCCPUBY	MEAN: All cpus percent busy (dispatched)	64.80	61.17	55.73	44.85	71.98	70.31	7
6	01DEC2001	PCCPUBY	MEAN: All cpus percent busy (dispatched)	60.52	55.36	49.98	38.51	63.16	66.37	6
7	01JAN2002	PCCPUBY	MEAN: All cpus percent busy (dispatched)	63.93	61.58	53.43	41.73	69.62	70.15	6
8	01FEB2002	PCCPUBY	MEAN: All cpus percent busy (dispatched)	64.96	60.55	58.71	45.40	67.43	76.15	6
9	01MAR2002	PCCPUBY	MEAN: All cpus percent busy (dispatched)	70.80	62.55	64.32	47.60	70.06	75.12	7
10	01APR2002	PCCPUBY	MEAN: All cpus percent busy (dispatched)	70.13	61.00	58.47	40.87	71.81	73.09	7

Figure 25. Transposed z/ OS Data

You are now ready to start TSFS, point it to this data set and start the analysis.

Figure 26 is the initial TSFS screen with the necessary parameters filled in. This window was activated by choosing the “Solutions” menu and then “Analysis” and finally “Time Series Forecasting System” from SAS on the PC.

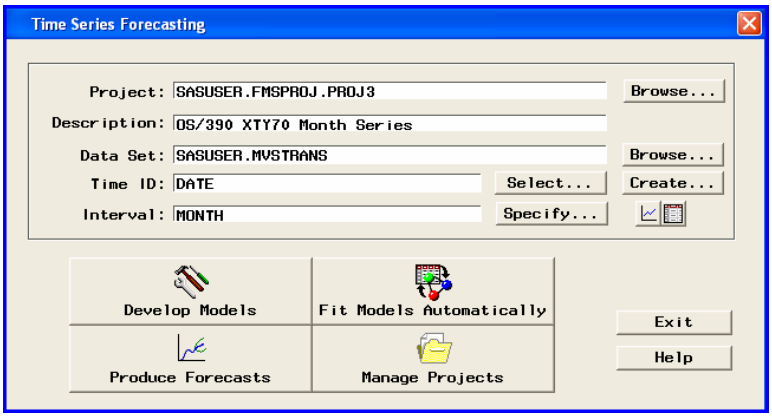


Figure 26. TSFS Selections

Figure 27 shows the results after you click the “Fit Models Automatically” button, which will find the best model for each time series in the data set that you are using. It will fit 24 time series; one time series for each hour.
The “Select…” button at the right of the “Selection Criterion” field (see Figure 27) enables the user to select parameters that determine how model selection will occur. For this example, the Root Mean Square Error criterion is selected.

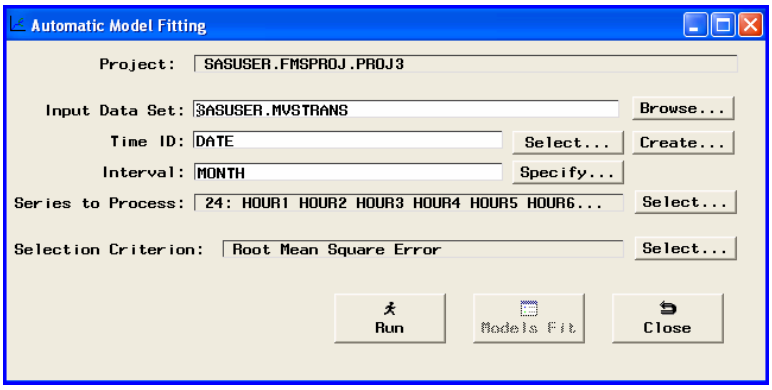
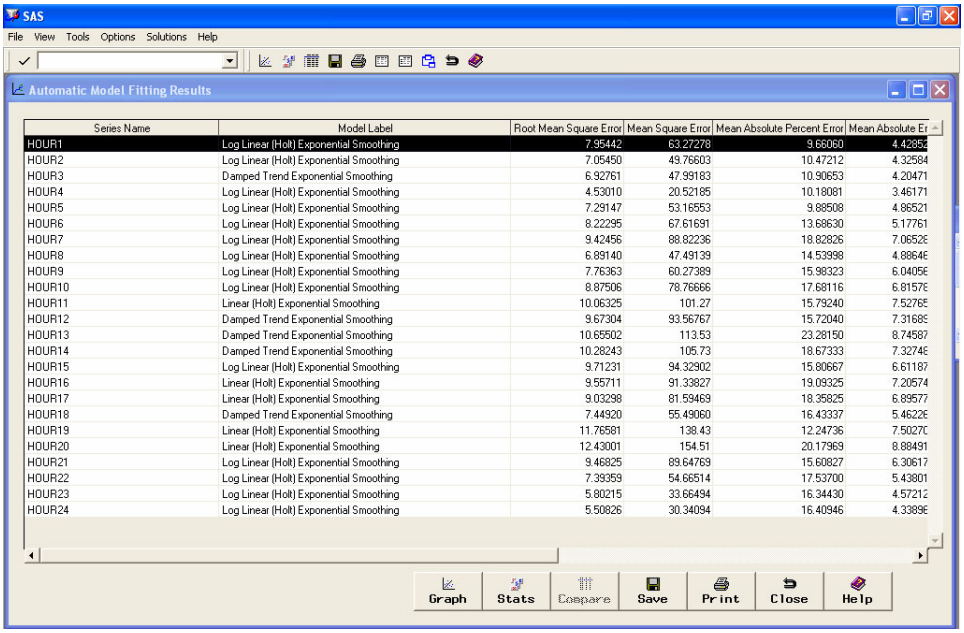


Figure 27. Fitting Models

Now, click the “Run” button, and TSFS will automatically find the best model for each time series and display the results, as shown in Figure 28.



Series Name	Model Label	Root Mean Square Error	Mean Square Error	Mean Absolute Percent Error	Mean Absolute Er
HOUR1	Log Linear (Holt) Exponential Smoothing	7.95442	63.27278	9.66060	4.42852
HOUR2	Log Linear (Holt) Exponential Smoothing	7.05450	49.76603	10.47212	4.32584
HOUR3	Damped Trend Exponential Smoothing	6.92761	47.99183	10.90653	4.20471
HOUR4	Log Linear (Holt) Exponential Smoothing	4.53010	20.52185	10.18081	3.46171
HOUR5	Log Linear (Holt) Exponential Smoothing	7.29147	53.16553	9.88508	4.86521
HOUR6	Log Linear (Holt) Exponential Smoothing	8.22295	67.61691	13.68630	5.17761
HOUR7	Log Linear (Holt) Exponential Smoothing	9.42456	88.82236	18.82826	7.06528
HOUR8	Log Linear (Holt) Exponential Smoothing	6.89140	47.49139	14.53998	4.88648
HOUR9	Log Linear (Holt) Exponential Smoothing	7.76363	60.27389	15.98323	6.04056
HOUR10	Log Linear (Holt) Exponential Smoothing	8.87506	78.76686	17.68116	6.81578
HOUR11	Linear (Holt) Exponential Smoothing	10.06325	101.27	15.79240	7.52765
HOUR12	Damped Trend Exponential Smoothing	9.67304	93.56767	15.72040	7.31685
HOUR13	Damped Trend Exponential Smoothing	10.65502	113.53	23.28150	8.74587
HOUR14	Damped Trend Exponential Smoothing	10.28243	105.73	18.67333	7.32748
HOUR15	Log Linear (Holt) Exponential Smoothing	9.71231	94.32902	15.80667	6.61187
HOUR16	Linear (Holt) Exponential Smoothing	9.55711	91.33827	19.09325	7.20574
HOUR17	Linear (Holt) Exponential Smoothing	9.03298	81.59469	18.35825	6.89577
HOUR18	Damped Trend Exponential Smoothing	7.44920	55.49060	16.43337	5.46228
HOUR19	Linear (Holt) Exponential Smoothing	11.76581	138.43	12.24736	7.50270
HOUR20	Linear (Holt) Exponential Smoothing	12.43001	154.51	20.17969	8.88491
HOUR21	Log Linear (Holt) Exponential Smoothing	9.46825	89.64769	15.60827	6.30617
HOUR22	Log Linear (Holt) Exponential Smoothing	7.39359	54.66514	17.53700	5.43801
HOUR23	Log Linear (Holt) Exponential Smoothing	5.80215	33.66494	16.34430	4.57212
HOUR24	Log Linear (Holt) Exponential Smoothing	5.50826	30.34094	16.40946	4.33896

Figure 28. Model Results for Time Series

Each time series model can be viewed by clicking the “Graph” button. Notice that the top right icon is pressed. The first-hour graph over the 2-year series, which shows the actual data value points and the prediction model line, can be seen in Figure 29.

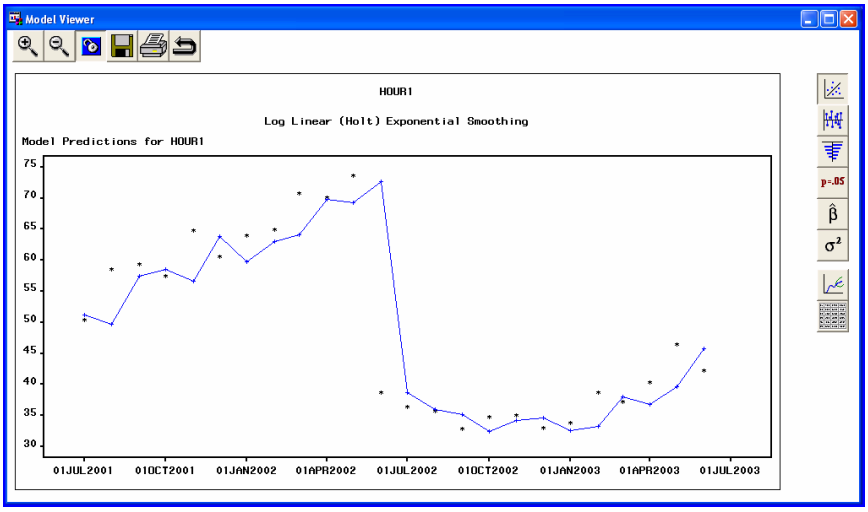


Figure 29. Viewing a Time Series

Notice that in June 2002 the processor usage dropped dramatically for this processor because a significant amount of work was off-loaded to another system. This indicates that this particular data sample should be run starting in June 2002. Because this paper is focused on how to use the interface, I will merely note this discrepancy and continue showing the concepts of using TSFS.

Using the icons on the right side of the Model Viewer window provides additional information about the specific series model. For more details about using these icons, see the SAS Course “Introduction to Time Series Forecasting Using SAS/ETS Software”. Let’s look at some of the icons now, starting with the second last icon shown in Figure 30 – the Forecast icon. This icon shows the prediction of HOUR1, two months into the future, based on the log linear exponential smoothing model.

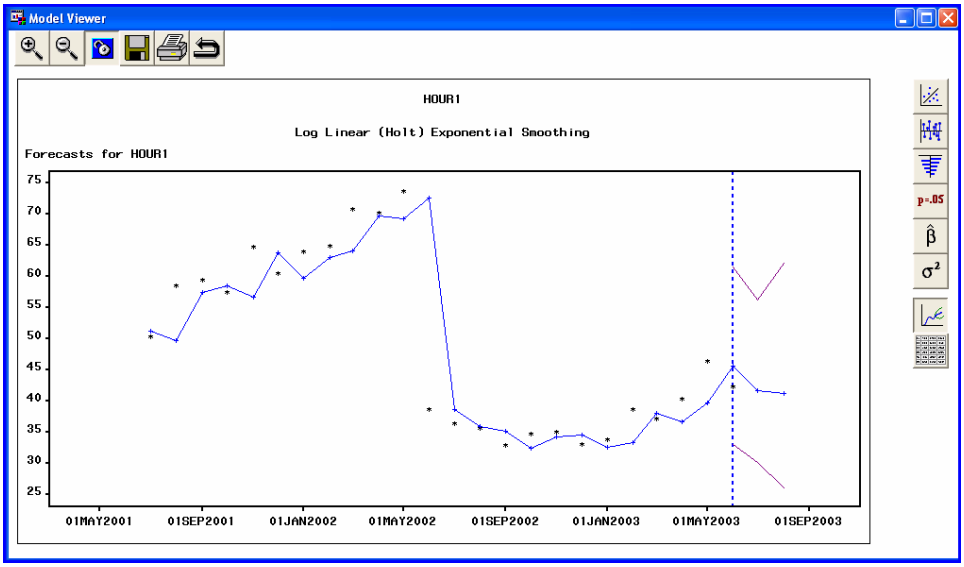


Figure 30. Viewing the Forecast

If you click the third icon from the top, the prediction error autocorrelation plots display.

By examining the left plot in Figure 31, you see from the random pattern within the two curved lines that reflect about 0 that there is no autocorrelation.

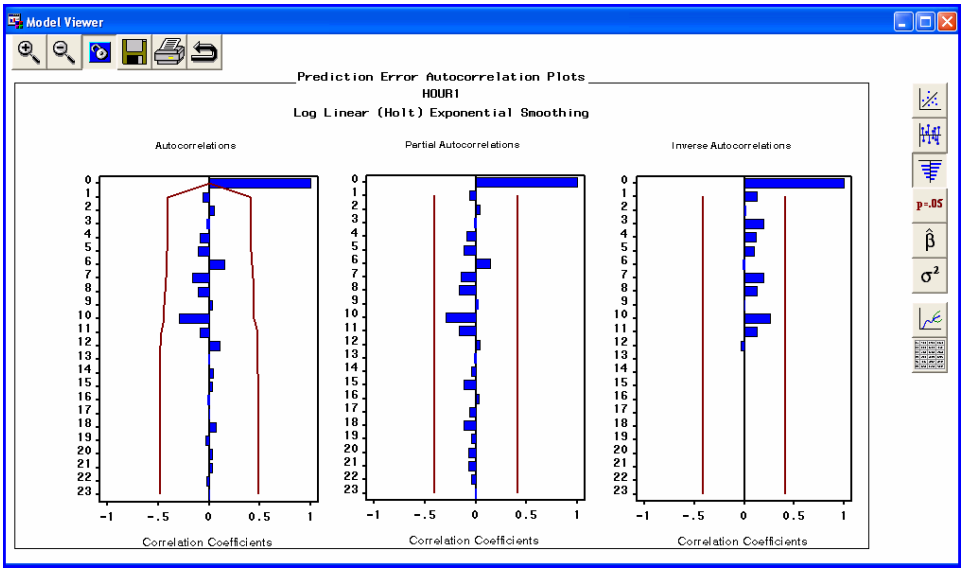


Figure 31. Autocorrelations

You can determine how significant your prediction model is by checking the statistics of fit in Figure 32. For an explanation of these factors, please refer to the TSFS Course notes.

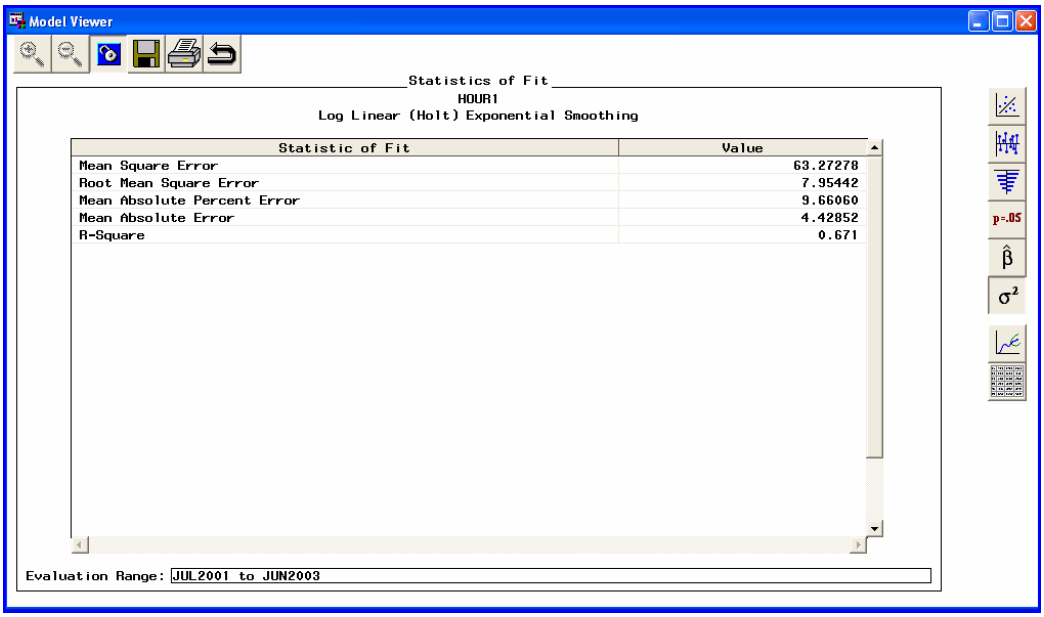


Figure 32. Statistics of Fit

Additional Considerations: Forecasting with a PDB

Rather than running a time series for each hour, an alternate approach is to create another monthly table in the PDB by using the Generic Collector facility of SAS IT Resource Management, and modifying the classification variables for the Month level of the table that you want to forecast. Changing the class variable list from 'MACHINE DATETIME HOUR SHIFT', to 'MACHINE DATETIME SHIFT' will remove the hourly granularity at the Month level, but will still allow you to keep statistics such as max, mean, and minimum for each month. With a table in this format, you can easily run a time series to predict the maximum value for specific metrics in your PDB, month to month.

Within TSFS, rather than fit models automatically, models can be developed specifically. Attending the training course will provide the analyst with the necessary skills to design appropriate models.

Conclusion

SAS Enterprise Guide makes analysis tasks easier and provides a drag-and-drop interface to perform complex SAS functions without requiring the user to write SAS code. SAS Enterprise Guide also generates SAS code that you can use to develop additional analyses that can be run automatically. Several examples of how to use SAS Enterprise Guide have been provided in this paper. The knowledge of how to perform these tasks will help you learn to solve business problems quickly with the help of SAS Enterprise Guide.

The Time Series Forecasting System will automatically fit models for time series data that can be used to help predict future requirements.

Why SAS?

In an era when many technology companies are promising you everything, SAS gives you the assurance that its Business Intelligence solutions are built on the proven technological capabilities of SAS software. For over 25 years, SAS has been on an unwavering path of strong, steady growth. As the world leader in Business Intelligence software and services, SAS is known for giving its customers the unmatched power and reliability of its solutions, unparalleled customer service and the ability to provide what our customers want and need most -- *The Power to Know.*

Founded in 1976, SAS is the only vendor that completely integrates leading data warehousing, analytics and traditional BI applications to create intelligence from massive amounts of data. Serving more than 38,000 businesses, government and university sites in 119 countries, our customers include 99 of the Fortune 100 and 90 percent of Fortune 500 companies.

For more information about SAS, visit www.sas.com.

World Headquarters
and SAS Americas
SAS Campus Drive
Cary, NC 27513 USA
Tel: (919) 677 8000
Fax: (919) 677 4444
U.S. & Canada sales:
(800) 727 0025

SAS International
PO Box 10 53 40
Neuenheimer Landstr. 28-30
D-69043 Heidelberg, Germany
Tel: (49) 6221 4160
Fax: (49) 6221 474850

www.sas.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.
Copyright © 2002, SAS Institute Inc. All rights reserved. **271821US.0304**