

SESUG 2022 Paper 209

A unique and innovative end-to-end demand planning and forecasting process using a collection of SAS products

Stephen Sloan; Data Science Senior Principal; Accenture
Kevin Gillette; Analytics Innovation Principal; Accenture Federal Services
Sushil B. Kumar; Application Development Senior Analyst; Accenture
Puspita Kumari Padhi; Software Engineering Team Lead; Accenture

ABSTRACT

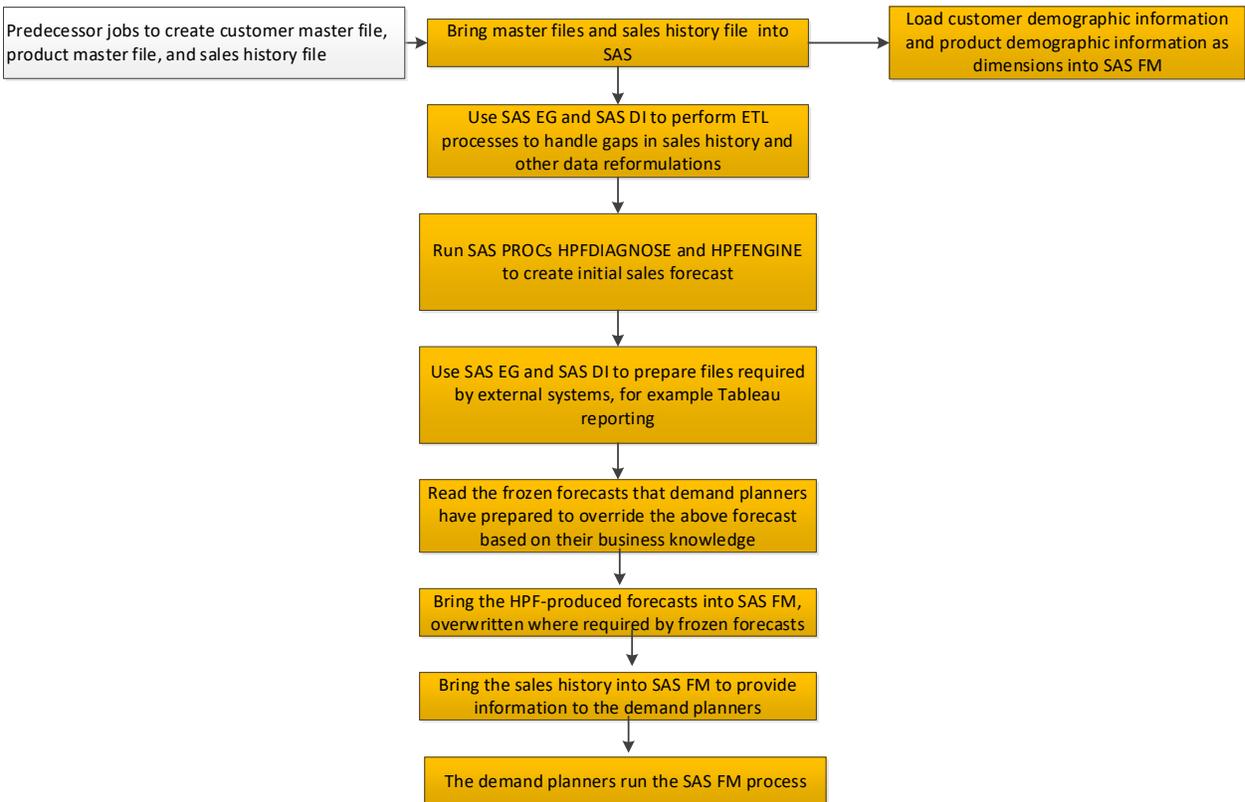
Forecasting demand can be a very tricky process. Questions arise about which statistical algorithms to use when forecasting based on past sales, how to incorporate business knowledge into the forecast, planning for unforeseen events, and planning for unique events that would not be predictable from sales history. As an example, COVID caused many forecasts to be wrong about quantities when purchasing switched from services to goods, from in-store sales to remote purchases, and from work in the office to remote work. Even when the total quantities of a product were forecasted incorrectly, the percentage distribution of the sales for some of the subcategories within the larger categories was often accurate, and vice versa. An example of this is the switch from eating and drinking in restaurants to take-out, where the total quantities of the items might have been the same, but the packaging changed (fewer kegs, more six-packs). The solution, then, is to leverage the useful information from the statistical forecasts while allowing the people who know the business to make individual or mass updates. All of this can be accomplished using existing SAS products: using SAS EG and SAS DI to read in, manipulate, and output data; using the High-Performing Forecasting (HPF) SAS PROCs, which underpin SAS Enterprise Miner, to create statistical forecasts; and using SAS Financial Management (SAS FM), which incorporates Excel features while remaining within SAS, to allow users to make individual and mass changes to SAS data sets.

INTRODUCTION

We outline a step-by-step process below that enables a combination of flexibility in testing and choosing statistical forecasts, allows knowledgeable businesspeople to adjust and test overrides to the forecasts, and tests and manipulates data throughout the process.

The two parts of the process are below. The first part describes bringing in the information from the source systems and preparing that information for use in the SAS FM product. The second part describes extracting the completed forecasts from SAS FM and preparing them for the supply planning system.

Part 1 – Get the source data, use the HPF PROCs and prepare it for SAS FM, and manipulate it using SAS FM.



STEP #1: READ AND PROCESS THE TRANSACTION HISTORY FILE AND THE MASTER FILES

This first step uses SAS DI and SAS EG to bring in the transaction history that will be used for the forecasting. This step also brings in any master files that can be used to expand on the codes contained in the transaction history files by getting all the associated demographic information.

This step uses SAS EG and SAS DI to check for anomalies and check for errors in the input files. Some errors could be:

- Out-of-range dates in the transaction files
- Unexpected or missing amounts in the transaction files
- Lack of uniqueness in the master files
- Referential integrity problems when performing lookups in the master files

At this point, some clean-up and re-extraction might be required. The clean-up is important because the statistical forecast can be erroneous if there are errors (even seemingly trivial errors) that cause certain time periods to have less or more volume than they really have. In addition, the HPF PROCs need to see the data in a specific, pre-defined format.

Finally, the input files might be in non-SAS databases like Oracle, Excel, or SAP, so we need to make sure that our SAS implementation has the necessary access capabilities (SAS ACCESS products).

STEP #2: CREATE THE STATISTICAL FORECAST AND ALLOW FOR USER UPDATES

SAS provides a wealth of forecasting tools for the developer and end-user alike. Our approach is to use the existing PROCs, principally PROC HPFDIAGNOSE and PROC HPFENGINE, in open code to permit maximum transparent customization - “fit for purpose” - with regard to the numerous forecasts. These two PROCs in particular work as a tandem unit – HPFDIAGNOSE examines an entire battery of possible forecasting models, and HPFENGINE helps to select the best model from among one or two best candidates.

The open-code approach is not simply to give the developers more code to write; it proves essential if there is a need to run forecasts in parallel for performance reasons. Most modern SAS servers are multi-CPU boxes, which means that the SAS administrator can configure SAS to take full advantage of the array of computing resources. Thus, if a forecasting problem admits a natural “parallelization” - a natural splitting of the forecasts being run into distinct cohorts that can run at the same time (no forecast is dependent on a prior forecast, for instance), then careful and judicious coding of the forecast logic can exploit this. This is not a technique that is do-able using any of the visual analytics tools in SAS (as far as this author knows); it is a combined effort of the SAS code and any script (a shell script, for example) that invokes it.

Moreover, forecasting is as much art as science, and frequently the forecaster or statistician can give the SAS tools the benefit of their insights into the data. An excellent case of this is the recent pandemic due to COVID-19 which has seriously disrupted sales and consumption patterns across nearly every part of the retail spectrum. Where the SAS automated tools may be fooled by deceptively bad data, the forecaster can step in. The system described here operates on an extensive series of default settings; however, the forecaster may override these settings for subsets of the data across multiple levels of stratification. Moreover, the system also permits the forecaster to re-run arbitrary subsets of the forecast bundle with (again) custom parameter settings. This includes model families under consideration, holdout sets for training and validating forecasts, statistical level for admission of parameters into the final forecast formula, and so on.

On the topic of data curation, the system described here also provides a number of algorithms to detect and correct data outliers. These algorithms include smoothing techniques (e.g., median filtering), back-filling, and “clipping” of outliers by use of windowed control limits. Future techniques might include change point analysis as well as other statistically robust heuristics. All of these curation techniques are written in open SAS code (DATA steps with a few judicious PROCs for sorting and reformatting).

STEP #3: CREATE ANY USEFUL REPORTS AND FILES FOR OTHER SYSTEMS

The current and previous information produced by the forecasting process, sometimes combined with information from other systems, can be valuable to other functions and systems, so this step takes these requirements into account by producing extract files and reports that are useful beyond the forecasting and planning system.

STEP #4: BRING IN THE FROZEN FORECASTS

As part of the SAS FM process, demand planners with the business knowledge that allows them to override statistical forecasts can use their knowledge to override the forecasts. Having a frozen forecast file makes it

easier for them to use saved forecasts in the future. Otherwise they need to modify the statistical forecast overrides every week to restore the forecasts to what they originally decided was appropriate.

STEP #5: PREPARE THE DATA FOR LOADING INTO SAS FM

The appendix contains a more complete description of the features and capabilities of SAS FM. It also includes screen shots that illustrate the process from the user's point of view. The forecasts coming out of the HPF PROCs, along with the frozen forecasts, will probably need to be modified before being entered into FM. Some possible modifications include:

- The record format, how the variables will be structured and positioned
- The level of summation, or granularity, of the forecasts
 - By product or product category
 - By geography
 - By type of customer
 - By time period
 - Daily
 - Weekly
 - Monthly
 - Quarterly
 - Annually
 - Other time periods that might be part of the organization's reporting structure
- Checks for reasonableness

SAS FM works with predefined dimensions, or "slicers", that need to be loaded into the database along with the forecasts so that mass changes can be made and tested within SAS FM. SAS FM is a very versatile product, combining the functionality of both SAS and Excel and providing a financial management process that can also be used for other purposes like forecasting. However, SAS FM is very specific about how to structure the data that it uses, so the dimension files need to be extracted and reformatted from the master files mentioned above and then imported into the SAS FM database.

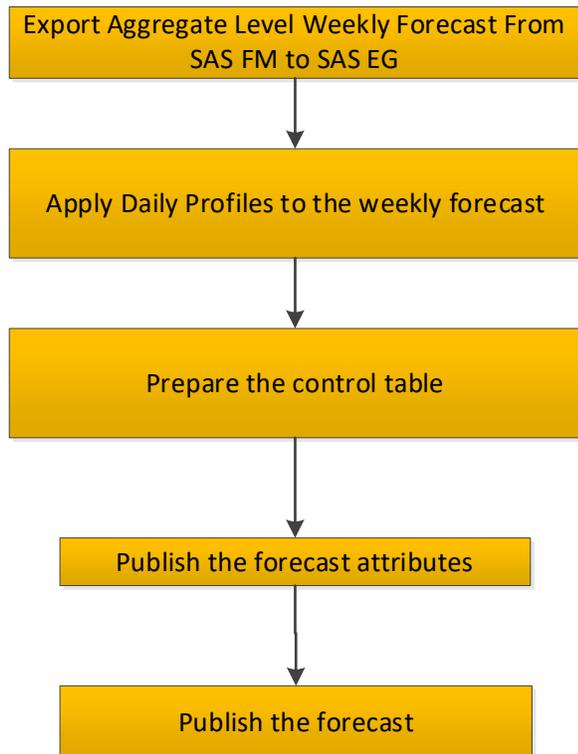
As mentioned in sections below, there are specific items that we might want to include in the SAS FM database that are independent of the statistical forecasts.

- Design preliminary forecasts for new products that have limited or no history so that they can be loaded into SAS FM. Our implementation achieved this by allowing the demand planners to give total forecasts and then allocating them to different subsets based on existing products that are similar.
- Pre-emptively exclude forecasts of data items scheduled for decommissioning or withdrawal.
- Enable product substitution and cannibalization.
- Enable the saving of modifications to the forecasts that occurred during SAS FM; allow the planners to ensure that certain forecasts don't get overwritten the next time the statistical forecast is run. This is the frozen forecast file mentioned above.

This is also a point at which files using current and previous forecasts and actual values could be created for output to other reporting and analytical systems.

STEP #6: USE SAS FM TO MODIFY THE FORECASTS AND ENSURE ALIGNMENT WITH EXTERNAL INFORMATION ABOUT SALES TRENDS AND ORGANIZATIONAL PLANS

Part 2 – Use the output from SAS FM, manipulate it based on your organization’s requirements.



STEP #1: EXPORT THE FORECASTS FROM SAS FM

Use the export feature in SAS FM to take the forecasts and bring them into SAS EG

STEP #2: APPLY THE DAILY PROFILES TO THE FORECAST

Our implementation used the HPF PROCs and SAS FM to create weekly forecast files. Then we used daily sales information to create daily profiles and apply them to the weekly forecasts.

STEP #3: MANIPULATE THE FORECASTS BASED ON A CONTROL TABLE THAT DEFINES ORGANIZATIONAL REQUIREMENTS

The forecasts coming out of SAS FM might need to be modified before being sent on to the downstream systems. As with the output from HPF the following items might need adjustment:

- The record format, how the variables will be structured and positioned
- The level of summation, or granularity, of the forecasts
 - By product or product category
 - By geography
 - By type of customer
 - By time period
 - Daily
 - Weekly
 - Monthly
 - Quarterly
 - Annually

- Other time periods that might be part of the organization’s reporting structure
- Checks for reasonableness

In addition, we would want to identify and save forecasts to be kept, or “frozen”, regardless of future statistical forecasts (see section below on saving forecasts).

Finally, the output files might need to be in non-SAS databases like Oracle, Excel, or SAP, so we need to make sure that our SAS implementation has the necessary access capabilities.

STEP #4: PUBLISH FORECAST ATTRIBUTES

Attributes like the type of forecast and the forecast version might be necessary for the supply planning system, so cross-reference tables are supplied so the downstream systems can look up the attributes.

STEP #5: PUBLISH THE FORECAST

This is the end goal of the process, creating a forecast and making it available to the supply planning system and other downstream systems.

OTHER ITEMS THAT CAN BE PART OF THE END-TO-END FORECASTING PROCESS

In addition to the steps in the sequential process itemized above, there are a few additional processes using SAS EG and SAS DI that can be incorporated into the end-to-end forecasting process.

Forecasting new products

When creating forecasts for an organization, it can be important to supply a complete forecast so that the organization can have a good idea about what components to order, what markets will be served, and how to allocate inputs. This means that new products whose sales are not planned to replace part or all of existing products will have to be included in the forecast, even though there is limited or no history to use to make the forecast. When making decisions about new products, the organization will generally have a good idea about how much will be sold over a certain period but might not have a good idea about how the sales will be allocated among different customers, locations, and subcategories or how quickly the sales will increase. Therefore, it will be necessary to develop a process to allow the people with product expertise to determine which products or categories are like the new products and how quickly the sales will scale up. The allocations mentioned above can then be applied to the new products.

Exclude items that are no longer sold or are scheduled to be withdrawn

The organization will have knowledge about products that either are no longer sold or are scheduled to be withdrawn, but these products might still receive forecasts from HPF because they have sales history during the time period used for forecasting. It’s not enough to check for the lack of recent sales because certain products could be seasonal and could thus legitimately have no recent sales but would still be live products and would require a forecast. What’s desirable is to allow the planners in the organization to identify products or customers or geographic areas or combinations or them for exclusion from the forecasts.

Forecasting product substitution and cannibalization

For the purpose of this paper, substitution is defined as a new product or group of products completely replacing another product or group of products, perhaps overall or perhaps within certain areas, while cannibalization is defined as a new product or group of products taking part of the sales from another

product or group of products, perhaps overall or perhaps within certain areas. For substitution, the planner will be able to identify a new product and the product that it's replacing so that HPF can use the replaced product's history to forecast for the new products. The process for cannibalization will be similar, except that the planner will identify what percentage of the sales of the cannibalized product will be lost to the new product. Refinements like projecting the possible increases in sales can be handled within SAS FM.

Retaining the forecasts produced by SAS FM instead of using the statistical forecasts produced by HPF

Resources within the organization often have knowledge of upcoming promotions, external trends in the marketplace, and other items that would not be picked up in a statistical forecast based on historical sales. SAS FM allows people to override the statistical forecasts produced by HPF and use their knowledge to modify the forecasts before the forecasts are sent downstream to the next department for use. However, the end-to-end process would use the statistical forecasts produced by HPF to override the previous FM forecasts during the next iteration of the process. To prevent the users from having to continuously modify the statistical forecasts, a method needs to be crafted to identify "frozen forecasts", thus allowing users to identify forecasts at different levels of granularity that need to be saved and not overridden by the statistical forecasts.

VALIDATION

The purpose of the HPF and SAS FM forecasts is to help the organization anticipate demand and make plans relating to purchasing components, staffing, distribution, marketing, and other areas. Therefore, it is important to be able to see how accurate the forecasts are on an ongoing basis. SAS EG can create projects with reports or output files that compare the forecasts to the actual sales on a regular basis and identify areas where the forecasts were accurate and areas where fine-tuning is necessary. The fine-tuning can be done while choosing different algorithms for HPF either at a universal or at a granular level and it can also be done within SAS FM as the discrepancies become apparent.

POINTS TO CONSIDER

SAS FM is very versatile and can handle a wide variety of dimensions, or "slicers". However, the underlying data within SAS FM is a Cartesian product of the different slicers and associated totals. Having too many slicers can degrade response time and occupy a large amount of space, so the SAS FM configuration needs to balance the desire for flexibility with the need for efficiency. Sometimes this can be resolved by having separate SAS FM forms for different groups of slicers that might be desired by different users, keeping in mind that the forecasts at the slicer level will need to be allocated to the product-customer combinations at a granular level.

Since the FM in SAS FM stands for Financial Management, we need to bear in mind that when we load an update into SAS FM, the results will show the difference between the previous and current values instead of the actual value being entered into a cell in SAS FM if that cell already has a non-zero value. This issue can be handled by either doing a full replacement of the previous FM forecast when the entire forecast is being loaded or by resetting specific items to zero before loading updates to those items.

The HPF PROCs provide a variety of options for creating the forecast and they do their best to choose the best forecasting algorithms. However, the people working with the forecasts often have better knowledge of what algorithms would work best. Features within HPF allow a statistician to override the option chosen

by HPF and use a more appropriate algorithm. Then HPF can be re-run using different algorithms until the statistician is satisfied with the algorithm chosen.

The number of years of actual history to use for the forecast and the number of years to forecast into the future present trade-offs to consider during implementation and operation. Ignoring other issues, more history will produce better forecasts. However, products might have changed, computer space might be at a premium, the environment might have changed so much that sales from a few years ago might be irrelevant, and other matters could lead to a desire for a shorter time horizon. Likewise, the more time that is forecast the more information is available, but there comes a point where it's not meaningful to forecast that far into the future. The amount of history and the length of the forecast need to be decided by the organization.

CONCLUSION

Sales forecasting is an inherently complicated process as it uses past history to project future trends and HPF provides a variety of methods for forecasting using past history (exponential smoothing, ARIMA, regression, etc.). In addition, there are trends in the external environment that might not be fully captured in past history, there can be exogenous events that were not predicted at all (e. g. COVID), and there are internally planned events like sales promotions that would not be accounted for in a statistical forecast. Using the power of HPF to produce forecasts and enabling statisticians to choose the forecasting algorithms, combined with SAS FM's enablement of the use of Excel within SAS at both granular and higher levels of data, mediated by programs written in SAS EG and SAS DI, allows an organization to create a comprehensive end-to-end forecasting process totally within the SAS environment.

APPENDIX

SAS Financial Management (SAS FM) is an advanced SAS solution for planning and reporting. Among other processes, it is designed to support the following management activities used in this application:

1. Data collection and retrieval
2. Management of dynamic hierarchical structures
3. Allocations and balancing entries
4. Reporting
5. Dynamic changes to forecasts, both individually and in groups, both with absolute amounts and percentages.

SAS FM User Interfaces:

There are three ways to interface with SAS FM:

SAS Financial Management Studio: A desktop application from which data administrators and other users build and manage the infrastructure that is required by the SAS FM Add-In for Microsoft Excel application and the SAS FM web application. For example, an administrator can use SAS Financial Management Studio to define the metadata that feeds into meaningful reports and forms.

SAS Financial Management Add-In for Microsoft Excel: SAS application that connects a desktop copy of Microsoft Excel to the SAS FM database. With this client application, business users can build and save specialized reports and forms. Other users (with permission) can use this application to view report data and enter data into forms for planning purposes. Data administrators also use the SAS Financial Management Add-In for Microsoft Excel client to create templates for planning.

SAS Financial Management Web Application: Web-based application that alerts the information consumer when a report has been published or alerts the information provider or information reviewer when an action is required. The SAS FM web application includes a workspace for managing and editing data-entry forms for financial planning.

Elements of SAS FM Data

Cycles and Dimension Types

In SAS FM, a cycle is a structured pool of stored data. There are a variety of dimension types that can be used in a cycle. In our application, since we were working with forecasts, we didn't use the more financial dimensions like currency and trader. These can be reworked into other types if necessary.

1. Account
2. Analysis
3. Currency
4. Frequency
5. Organization
6. Time
7. Source
8. Trader (a mirror of Organization)

Note: If necessary, Additional dimension types can be created to meet the requirements of the specific application.

Models A model is a structure for viewing and interacting with the data in a cycle. The model is the basis of both forms and reports. The structure of a model includes a set of hierarchies (from

required and optional dimension types) as well as rates and formulas. A model can also have its own set of adjustments and rules.

SAS Financial Management Server Configuration

1. Metadata server—Server on which the SAS Metadata Server software is running. SAS must be available on this same machine.
2. Data Tier server—Server on which SAS runs data-handling programs (including the logical servers for SAS Workspace and SAS Stored Process servers). Transformations, error tables, and jobs are installed on the data tier server. Note: The same machine is often used as both the data tier server and the metadata server.
3. Middle-Tier server—Server on which the SAS managed servers run.

Supplying Data to SAS Financial Management

1. Loading data:

Data will be loaded from its source, through the SAS FM staging area in SAS Data Integration Studio, to the SAS FM Data Mart. The Data Mart consists of PostgreSQL tables specific to SAS FM that support dimensional data, fact data, and special data. The staging area has a SAS library named StageFM. The StageFM library consists of approximately 85 SAS FM tables that are categorized as follows: dimension tables, fact tables, reference tables, and special tables. Typically, you load data from the StageFM library into the SAS FM Data Mart.

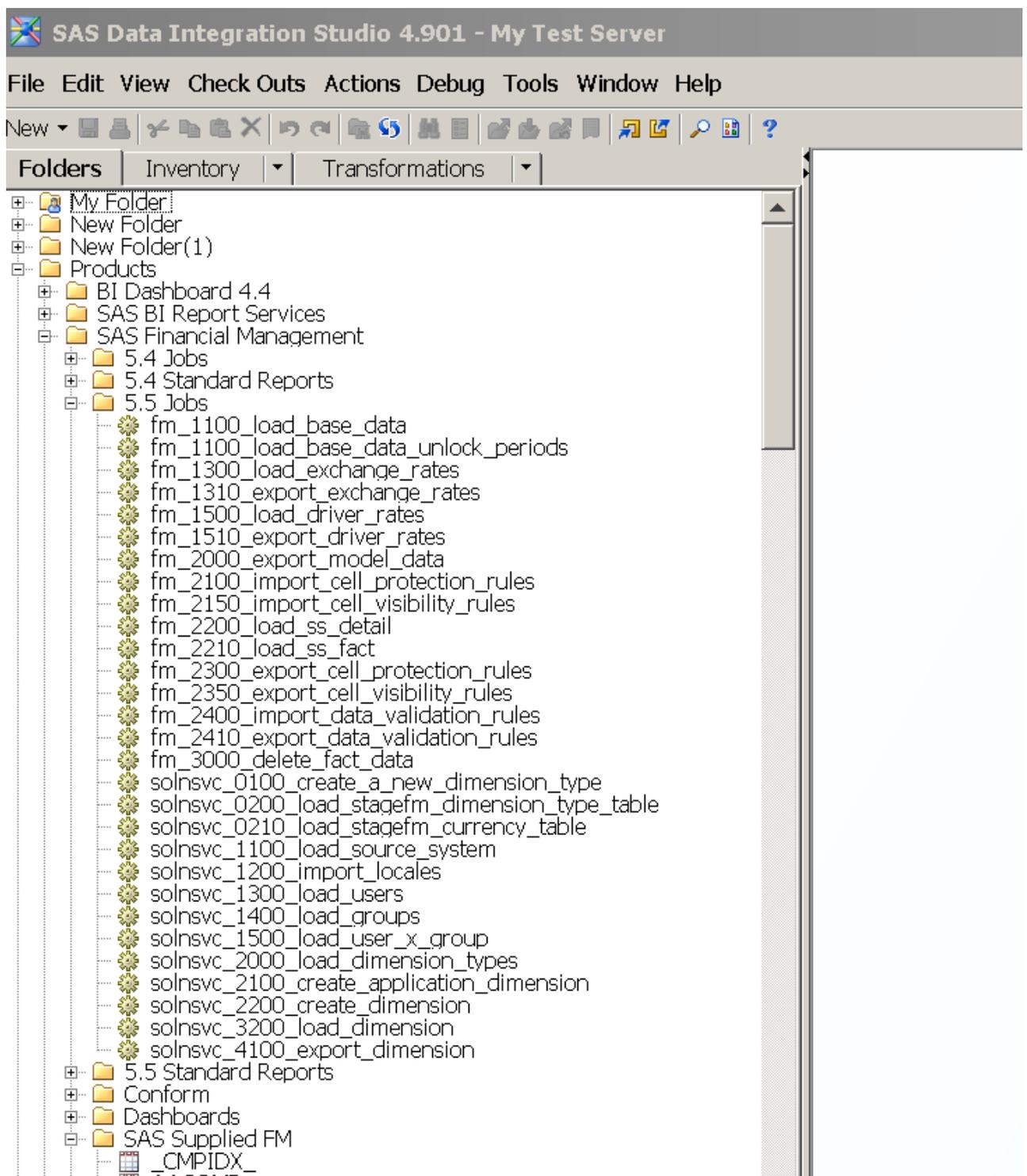
2. Loading data into staging tables:

If the data that you are loading is from a source outside of the SAS environment, you must write a job that extracts the data from its source and loads the data into the appropriate SAS FM staging table in SAS Data Integration Studio.

If the data that you are loading is from a SAS table of predefined data, you can use a SAS Data Integration Studio job to load the corresponding staging area table.

3. Loading data into the Data Mart:

There are specific DI jobs to load the data into data-mart. For specific types the data can be loaded from SAS FM's load wizard. Below is an illustration of a SAS FM process using SAS DI to extract the data.



3. Extending the staging area

For a staging area where the additional data can't be loaded into a predefined Data Mart, custom dimension types can be added whose members can be used to qualify the data for SAS FM. Each custom dimension type is supported by a set of four additional staging area tables. The data in these additional tables can be loaded into the Data Mart in the same way as data for the basic dimension types.

It exports accounting data/model data. To export data by using the fm_2000_export_model_data job, the following steps need to be done.

In SAS Data Integration Studio, select the Folders tab.

In the Folders tree, select Products -SAS Financial Management -5.5 Jobs.

Make a copy of the fm_2000_export_model_data job

Double-click fm_2000_export_model_data in the list of jobs. The job is displayed in the Job Editor window.

Right-click the export_model_data transformation and select Properties from the pop-up menu. The export_model data Properties window is displayed.

Select the Options tab.

Enter values for the following options: Option Description Result Code Code of the model that is the source of the data Table holding Period and Analysis Codes Name of a SAS data set that specifies the member combinations to export data for. The Precode region on the Precode and Postcode tab contains sample code that builds a SAS data set with the required layout.

3. Create and use a SAS macro.

The exported data from FM can be used to feed downstream systems.